LSHSS

Tutorial

The Complexity Approach to Phonological Treatment: How to Select Treatment Targets

Holly L. Storkel[a]

Purpose: There are a number of evidence-based treatments for preschool children with phonological disorders (Baker & McLeod, 2011). However, a recent survey by Brumbaugh and Smit (2013) suggests that speech-language pathologists are not equally familiar with all evidence-based treatment alternatives, particularly the complexity approach. The goal of this clinical tutorial is to provide coaching on the implementation of the complexity approach in clinical practice, focusing on treatment target selection.
Method: Evidence related to selecting targets for treatment based on characteristics of the targets (i.e., developmental norms, implicational universals) and characteristics of children's knowledge of the targets (i.e., accuracy, stimulability) is reviewed. Free resources are provided to aid clinicians in assessing accuracy and stimulability of singletons and clusters. Use of treatment target selection and generalization prediction worksheets is illustrated with 3 preschool children.
Results: Clinicians can integrate multiple pieces of information to select complex targets and successfully apply the complexity approach to their own clinical practice.
Conclusion: Incorporating the complexity approach into clinical practice will expand the range of evidence-based treatment options that clinicians can use when treating preschool children with phonological disorders.
Supplemental Material S1: https://doi.org/10.23641/asha.6007562
KU ScholarWorks Supplemental Material: http://hdl.handle.net/1808/24767

Children with speech sound disorders have trouble learning the sounds of their native language, often requiring clinical treatment to normalize their speech. Although there are a variety of speech sound disorders, this article will focus on children with functional phonological disorders. Children with functional phonological disorders exhibit delays in sound production in the absence of any obvious motoric, structural, sensory, cognitive, or neurologic cause. There are a number of evidence-based treatments for children with phonological disorders (see Baker & McLeod, 2011, for a review). However, a recent survey by Brumbaugh and Smit (2013) suggests that speech-language pathologists (SLPs) are not equally familiar with all evidence-based treatment alternatives. Specifically, Brumbaugh and Smit's survey showed that over 50% of SLPs sometimes, often, or always used a traditional, phonological awareness, minimal pairs, or cycles approach in treatment of children with phonological disorders. In contrast, only 8% of SLPs sometimes, often, or always used a complexity approach in treatment of children with phonological disorders. This is surprising given that the evidence base for the complexity approach is quite strong with "more research studies investigating the complexity approach than almost all other approaches combined" (Kamhi, 2006, p. 275). Why is there a lack of implementation of the complexity approach compared with other approaches? It is likely that clinicians lack familiarity with the complexity approach. In fact, Brumbaugh and Smit showed that 70% of SLPs were not familiar with the complexity approach. In addition, the complexity approach requires a detailed analysis of phonology to guide treatment planning, which may seem challenging to clinicians with high caseloads, but with appropriate support, the planning process can be streamlined. This clinical tutorial will focus on one aspect of the complexity approach: selecting complex treatment targets. The goals of this tutorial are to (a) review the evidence on the complexity approach to show that the initial investment in phonological analysis pays off in greater gains during treatment and (b) provide coaching

[a]Speech-Language-Hearing: Sciences & Disorders, University of Kansas, Lawrence

Correspondence to Holly L. Storkel: hstorkel@ku.edu

Editor-in-Chief: Shelley Gray
Editor: Ignatius Nip

Received August 10, 2017
Revision received October 12, 2017
Accepted October 29, 2017
https://doi.org/10.1044/2017_LSHSS-17-0082

Disclosure: The author has declared that no competing interests existed at the time of publication.

and implementation resources to speed the planning process in selecting complex treatment targets. Other resources can be used to learn about additional methods consistent with the complexity approach, such as complexity-based minimal pair treatment variants (e.g., maximal opposition, empty set; see Barlow & Gierut, 2002, for a summary).

In research studies, the complexity approach has typically been tested with 3- to 6-year-old children with very low scores (i.e., below the fifth or sixth percentile) on an articulation test (e.g., Goldman-Fristoe Test of Articulation) and at least five to seven target English sounds excluded from their phonemic inventories. The goal of clinical treatment for children like this who have multiple sound errors is to produce system-wide change in a child's speech so that acquisition is accelerated to close the gap with typically developing peers. System-wide change has been defined as multiple forms of generalization (Gierut, 1998, 2001). Specifically, the child produces the treatment target (e.g., liquid /r/) accurately in untreated words, produces untreated sounds related to the treatment target (e.g., liquid /l/) accurately in untreated words, and produces untreated sounds unrelated to the treatment target (e.g., fricative /θ/) accurately in untreated words. Thus, with system-wide change, the child learns something more global about the sound system of his or her native language with the treatment target serving merely as a trigger to induce broad phonological learning. When system-wide change occurs, a child improves overall intelligibility, rather than only improving a single or few treatment targets. Thus, a child should require less time in phonological treatment, leading to earlier dismissal from an SLP's caseload. System-wide change is the primary goal within the complexity approach (Gierut, 1998, 2001). Consequently, the lack of implementation of this evidence-based approach is a cause for concern.

Moreover, early system-wide change is a crucial target for speech normalization. Research in infant speech perception shows that children are quickly tuning their phonological system to the characteristics of their native language (see Werker & Hensch, 2015, for a recent review). In addition, studies of second-language learning show that older children and adults have phonological systems that are more solidified with less flexibility, making it difficult for second-language learners to master native-like pronunciation of speech sounds (Piske, MacKay, & Flege, 2001). These two bodies of research suggest that the plasticity of the phonological system decreases across development, even in the typically developing population. Within children with phonological disorders, Shriberg, Gruber, and Kwiatkowski (1994) demonstrated that there are periods of accelerated phonological learning followed by plateaus in phonological learning for children with phonological disorders. Specifically, there is a period of accelerated change from 4 to 6 years old, followed by a plateau from 6 to 7 years old. There is another period of accelerated change from 7 to 8.5 years old, followed by a final stable plateau beginning at 8.5 years old. Thus, it may be easier to make system-wide change in younger children (i.e., age of 4–6 years) and more difficult to induce change in older children (i.e., age of 8.5 years or older). Consequently, it is imperative to provide high-quality phonological treatment early so that children's speech has normalized before they reach a natural plateau, especially the final plateau beginning at 8.5 years old.

Taken together, the use of the complexity approach in phonological treatment has potential implications for caseload management. By targeting system-wide change via the complexity approach during a natural period of accelerated learning (i.e., preschool), there is greater potential to normalize speech sound development and transition children off caseloads before school entry. Within a complexity approach, targets are selected for treatment based on characteristics of the targets (e.g., developmental norms, implicational universals) and characteristics of children's knowledge of the targets (e.g., accuracy, stimulability).

Characteristics of Targets: Developmental Norms

Developmental norms document the age when most children produce a target sound accurately, with typical cutoffs being the age when 75% of children or more produce the target accurately (Templin, 1957) or the age when 90% of children or more produce the target accurately (Smit, Hand, Freilinger, Bernthal, & Bird, 1990). The treatment efficacy of selecting early- versus late-acquired targets was first examined by Gierut, Morrisette, Hughes, and Rowland (1996) in a pair of single-subject studies involving nine children (age = 3;5–5;6 years;months). In Study 2, which used a multiple-baseline approach with six children, early-acquired targets were defined as those that were acquired 1 year before the child's current chronological age, based on the norms of Smit and colleagues (1990). In contrast, late-acquired targets were those that were typically acquired 1 year or more beyond the child's chronological age, based on the Smit norms. For example, for a male child who is 4;6, targets with an age-of-acquisition of 3;6 or younger would be considered early acquired (/m n h w p b t d k/ and /f/ in initial position). For the same child, late-acquired targets would be those with an age-of-acquisition of 5;6 or later: /v, θ, ð, s, z, ʃ, ʧ, ʤ, l, r/ or /f/ in final position. Children were taught one target through imitation and spontaneous production of nonwords for a maximum of nineteen 1-hr sessions. In Study 2 by Gierut and colleagues, children taught early-acquired targets (i.e., /k, g, f/) and those taught late-acquired targets (i.e., /r, θ, s/) learned the treatment target and generalized it to untreated words. Likewise, both groups of children made change in sounds that shared manner with the treatment target. The main outcome that differentiated the groups is that children taught early-acquired targets made minimal change in untreated sounds that were unrelated to the treatment target. Specifically, unrelated untreated sounds were produced with 0%–10% accuracy after treatment. In contrast, children taught late-acquired targets demonstrated 30%–50% accuracy producing sounds unrelated to the treatment target (where these sounds were produced with 0% accuracy before treatment). This finding was replicated by Gierut and Morrisette (2012) who showed that 10 preschool children (age = 3;10–5;11) demonstrated greater system-wide change with treatment of a late-8 sound (/ʃ/) than with treatment of a mid-8 sound (/k f ʧ/). Taken together, treatment of late-acquired targets promoted greater system-wide change than treatment of early-acquired targets.

A follow-up randomized controlled group study by Rvachew and Nowak (2001) seemed to contradict this finding. Specifically, 48 preschool children received 12 weeks of once-per-week 30- to 40-min treatment sessions on four early-acquired or four late-acquired targets (i.e., two treatment targets per 6-week treatment block). However, age-of-acquisition was coupled with the child's knowledge of the treatment target. Thus, children were taught either most-knowledge/early-acquired targets or least-knowledge/late-acquired targets. Rvachew and Nowak showed that children taught least-knowledge/late-acquired targets completed fewer steps in treatment than children taught most-knowledge/early-acquired targets. Specifically, the highest treatment step achieved in six treatment sessions, on average, by children taught least-knowledge/late-acquired targets was 2.83, which corresponds to word level practice (i.e., Step 2 = imitated words, Step 3 = spontaneous words). In contrast, the highest treatment step achieved in six treatment sessions, on average, by children taught most-knowledge/early-acquired targets was 4.7, which corresponds to sentence level practice (i.e., Step 4 = imitated patterned sentences, Step 5 = spontaneous patterned sentences). In terms of change in accuracy of the treatment target, children taught least-knowledge/late-acquired targets showed lower accuracy in producing the treatment target in untreated words than children taught most-knowledge/early-acquired targets. However, there were no significant differences between the groups in overall accuracy of sound production. Thus, treatment of more-knowledge/early-acquired targets appeared to result in faster completion of treatment steps and greater learning of the treatment target, when only minimal treatment (i.e., six sessions) was provided.

How can Rvachew and Nowak's (2001) results be reconciled with those of Gierut? There are several differences across studies. First, Rvachew and Nowak combined developmental norms with knowledge, whereas Gierut and colleagues (1996) held knowledge constant at least knowledge and examined developmental norms in isolation. Second, Rvachew and Nowak held time constant at six sessions but allowed treatment steps to vary, whereas Gierut and colleagues had children in both conditions complete the same treatment steps. Thus, the two studies contrast different conditions. Rvachew and Nowak contrasted most-knowledge/early-acquired targets with treatment through the sentence level against least-knowledge/late-acquired targets with treatment through the word level, whereas Gierut and colleagues contrasted early-acquired targets with late-acquired targets while holding knowledge constant (i.e., least knowledge) and treatment constant (i.e., treatment at the word level). Because different conditions are contrasted, the findings from Rvachew and Nowak and those from Gierut and colleagues are complementary rather than contradictory. Rvachew and Nowak show that treatment of most-knowledge/early-acquired targets results in faster progress through treatment steps than treatment of least-knowledge/late-acquired targets. Moreover, treatment of most-knowledge/early-acquired targets with treatment through the sentence level may lead to better learning of the treatment target than treatment of least-knowledge/late-acquired targets through the word level, but this does not translate into differences in broad system-wide generalization. Gierut and colleagues show that treatment of least-knowledge/early-acquired targets and least-knowledge/late-acquired targets does produce differential sound learning when the same treatment steps are completed. Thus, if rapid completion of treatment steps in a short number of sessions or learning of the treatment target primarily is the goal of treatment, then treatment of most-knowledge/early-acquired targets through the sentence level may be optimal. On the other hand, if system-wide phonological change is the goal of treatment, then treatment of least-knowledge/late-acquired targets through a required set of treatment steps may be optimal.

Characteristics of Targets: Implicational Universals

Consonant Singletons

A singleton is a sound that occupies a syllable position in isolation (e.g., "bake" contains two singleton consonants: /b/ in the syllable onset and /k/ in the syllable coda). Implicational universals describe patterns that are observed across the world's languages and across individual speakers learning a language. In the case of phonology, implicational universals describe patterns of co-occurrences of sounds (Gierut, 2007). For example, one observation is that "if a language has fricatives, then it will also have stops." This is based on the observed patterns that languages can have (a) neither stops nor fricatives, (b) stops only, or (c) stops and fricatives. However, a language with only fricatives and no stops has not been observed. In this case, the sound class that can occur alone (e.g., stops) is referred to as unmarked, whereas the sound class that cannot occur alone (e.g., fricatives) is referred to as marked. The unmarked sound is assumed to be less complex (both phonologically and motorically), and the marked sound is assumed to be more complex (both phonologically and motorically). Table 1 shows the marked and unmarked classes for singletons (and clusters).

A variety of different single-subject studies have tested treatment of different marked and unmarked structures. For example, one study of eight children with hearing impairment (McReynolds & Jetzke, 1986) showed that children who were taught a voiced stop (i.e., /d/ or /g/) made greater change in accuracy of cognate stops (i.e., voiceless stop /t/ or /k/) than children who were taught a voiceless stop (i.e., /t/ or /k/). This suggests that treatment of

Table 1. Marked and unmarked sound classes based on implicational universals.

Marked (more complex) If a language has…	Unmarked (less complex) Then it will also have…
Fricatives	Stops
Affricates	Fricatives
Voiced obstruents (affricates, fricatives, stops)	Voiceless obstruents (affricates, fricatives, stops)
Liquids	Nasals
True clusters	Affricates
Small-sonority-difference clusters	Large-sonority-difference clusters
True clusters	Adjunct clusters
Three-element clusters	Two-element clusters

the marked (i.e.., more complex) voiced obstruent leads to greater change in the unmarked (i.e., less complex) voiceless obstruent than vice versa. A study of children with phonological disorders (Dinnsen & Elbert, 1984) showed that treatment of marked fricatives enhanced learning of unmarked stops but the opposite approach, treatment of unmarked stops, did not enhance learning of marked fricatives. A study of the phonemic inventories of 30 children (age = 3;4–5;7) with phonological disorders (Gierut, Simmerman, & Neumann, 1994) confirmed two implicational universals for children with phonological disorders: (a) the inventories contained affricates and fricatives, or fricatives only, but never affricates without fricatives, and (b) the inventories contained liquids and nasals, or nasals only, but never liquids without nasals. This finding confirms that affricates and liquids are more complex (i.e., marked), whereas fricatives and nasals are less complex (i.e., unmarked). Although specific implicational universals (e.g., fricatives imply stops) have only been tested typically in one single-subject study, the results across studies demonstrate a consistent pattern: Treatment of more complex, marked targets leads to greater system-wide change than treatment of less complex, unmarked targets.

Consonant Clusters

There also are relevant implicational universals for consonant clusters: one or more sounds that occupy a syllable position in tandem (e.g., “brake” contains one cluster /br/ in the syllable onset). First, clusters are more marked than singletons. Thus, languages have singletons only or have singletons and clusters, but a language with only clusters in the absence of singletons is unattested. Moreover, there appears to be a relationship between clusters and affricates such that languages have affricates only or clusters and affricates, but not clusters alone. Gierut and O’Connor (2002) examined the phonemic inventories of 110 children (age = 3;0–8;6) with phonological disorders and found that 94% (103/110) of inventories matched this implicational universal. Moreover, Gierut and Champion (2001) demonstrated that treatment of a cluster leads to widespread change in singletons, including affricates, further supporting the implicational universal.

Importantly, within various possible clusters, some clusters are more marked than others. In this case, markedness relates to the sonority sequencing principle (see Gierut, 1999, for a review). Sonority refers to the resonance of a sound. According to the sonority sequencing principle, sonority rises in the onset of a syllable, peaks at the nucleus (typically a vowel), and then falls in the coda. This leads to the expectation that an onset cluster will have rising sonority. The sonority of consonants is ranked as follows from least to most sonorous: voiceless stops/affricates, voiced stops/affricates, voiceless fricatives, voiced fricatives, nasals, liquids, and glides. In addition, arbitrary numbers can be assigned to this sonority ranking so that the difference in sonority within a cluster can be calculated: voiceless stops/affricates (7), voiced stops/affricates (6), voiceless fricatives (5), voiced fricatives (4), nasals (3), liquids (2), and glides (1). Thus, for the cluster /br/ in “brake,” the sonority difference is 6 (/b/, voiced stop) minus 2 (/r/, liquid) equals +4. The positive sign indicates that the sonority is rising, as expected. Note that the terms “sonority difference” and “sonority distance” are synonymous and used interchangeably in the literature. Both larger sonority differences (e.g., /kw/ = 7 − 1 = +6) and smaller sonority differences (e.g., /sm/ = 5 − 3 = +2) are observed in English. A full list of English onset clusters by sonority difference is shown in Table 2. In terms of implicational universals, clusters with larger sonority differences (e.g., /kw/, +6) are considered less marked than clusters with smaller sonority differences (e.g., /sm/, +2).

Gierut (1999) provides evidence that treatment of marked clusters leads to greater system-wide change than treatment of unmarked clusters. Specifically, three children (age = 3;8–7;8) were taught marked clusters with a small sonority difference (i.e., a difference of 3 or 4, /fl bl/), and three children (age = 3;2–6;10) were taught less marked clusters with a large sonority difference (i.e., a difference of 5 or 6, /kl pr kw/). Children who were taught less marked clusters with a large sonority difference learned their treated cluster and showed narrow generalization to other clusters that were superficially similar to the treated cluster (i.e., shared one sound in the cluster). For example, one child was taught /pr/, with a sonority difference of 5, and learned other *r*-clusters, such as /tr dr/. In contrast, children taught marked clusters with a low sonority difference showed broad system-wide change, learning their taught cluster (e.g., /fl/), other clusters superficially similar to the taught cluster (e.g., other *l*-clusters), and seemingly unrelated clusters (e.g., *r*-clusters). Moreover, examination of learned clusters by sonority difference showed that children taught less marked clusters with a large sonority difference (e.g., +5) learned clusters around that same sonority difference (e.g., +4, +5). In contrast, children taught marked clusters with a small sonority difference (e.g., +3) learned other clusters at a variety of sonority differences (e.g., −2, +2, +3, +4, +5, +6). Taken together, broad system-wide change was observed when children were taught marked clusters with a small sonority difference (i.e., a difference of +3 or +4). A similar pattern was observed in a study contrasting

Table 2. Sonority difference for onset clusters.

Sonority difference	Onset cluster elements	Onset cluster examples
6	Voiceless stop + glide	/tw/, /kw/, /pj/, /kj/
5	Voiced stop + glide	/bj/
	Voiceless stop + liquid	/pl/, /kl/, /pr/, /tr/, /kr/
4	Voiced stop + liquid	/bl/, /gl/, /br/, /dr/, /gr/
	Voiceless fricative + glide	/sw/, /fj/
3	Voiced fricative + glide	/vj/
	Voiceless fricative + liquid	/fl/, /sl/, /fr/, /θr/, /ʃr/
2	/s/ + nasal	/sm/, /sn/
	Nasal + glide	/mj/
−2	/s/ + stop	/sp/, /st/, /sk/

treatment of more marked fricative + liquid clusters and less marked stop + liquid clusters for six children (age = 4;4–6;3; Elbert, Dinnsen, & Powell, 1984; Powell & Elbert, 1984).

To this point, I have only considered two-element true clusters in English. English also has adjunct clusters (i.e., /s/ + stop). Adjunct clusters violate the sonority sequencing principle (Gierut, 1999). The fricative /s/ has a sonority of 5, and the voiceless stops have a sonority of 7. Thus, the sonority difference for /s/ + stop clusters is −2 (i.e., 5 − 7 = −2), indicating that the sonority is falling, which is unexpected in the onset. In Gierut's (1999) treatment study, it appeared that treatment of adjunct clusters facilitated learning of adjunct clusters with limited generalization, primarily to clusters that were superficially similar to the treated cluster (i.e., clusters containing /s/). In contrast, as previously noted, treatment of true clusters led to change in true clusters and adjunct clusters (Gierut, 1999). In this way, adjunct clusters are considered less marked than true clusters. Thus, treatment of true clusters with a small sonority difference is prioritized over adjunct clusters within the complexity approach. There also was some indication in Gierut's (1999) study that children may group /s/ + nasal clusters (/sm sn/) with a sonority difference of 2 with /s/ + stop adjunct clusters with a sonority difference of −2 because of superficial similarity (i.e., all are *s*-clusters). Specifically, when taught adjunct clusters, children improved production of other *s*-clusters, which raises the possibility that children may think of all *s*-clusters as having an adjunct structure. The implication for treatment is that targeting /s/ + nasal clusters may not lead to system-wide change, although /s/ + nasal clusters have a small sonority difference (+2). More research is needed to fully understand this relationship among true *s*-clusters and adjunct *s*-clusters. Thus, although /s/ + nasal clusters have the smallest sonority difference, I encourage focusing on treatment of true clusters with sonority differences of 3 and 4 to closely match the conditions where Gierut observed the broadest learning.

English also has three-element clusters. Three-element clusters in English (/skw spl spr str skr/) are interesting because they contain both an adjunct and a true cluster. All three-element clusters in English begin with /s/ and a voiceless stop, which is the adjunct structure that violates the sonority sequencing principle with a sonority difference of −2. The third element is always a glide or liquid, where voiceless stop and glide/liquid correspond to the structure of a true cluster with a sonority difference of 5 or 6. Perhaps, teaching a three-element cluster would spark learning of both adjunct and true clusters as well as singletons. Gierut and Champion (2001) tested this possibility with eight children (age = 3;4–6;3) with phonological disorders. This is the one study within the complexity approach that suggests that children may need a foundation for learning the most complex targets of the language. That is, what children learned from being taught a three-element cluster depended on their knowledge of the phonological system at pretreatment. In this case, knowledge was based on a phonemic inventory of singleton consonants. Children who had the /s/ and stop of their three-element clusters in their phonemic inventory seemed to focus more on the adjunct element of the three-element cluster, demonstrating learning of untaught two-element adjunct clusters but nothing else. In contrast, children who had the stop and glide/liquid of their treated three-element cluster in their phonemic inventory seemed to focus more on the true cluster element of the three-element cluster, showing learning of untaught two-element true clusters and singletons. Finally, children who knew only one element of their treated three-element cluster did not show generalization to any two-element structures (i.e., adjuncts or true clusters). Taken together, selection of a three-element cluster as a treatment target requires careful consideration because it is the most advanced phonological structure in English. Specifically, it is recommended that a three-element cluster be selected as a treatment target only if the child "knows" the target stop and glide or liquid as a singleton to focus learning on the true cluster element of the three-element cluster to induce broad system-wide change (Gierut & Champion, 2001). Interestingly, there is no evidence that other complex targets (i.e., two-element clusters, singletons) require this type of foundational knowledge for children to benefit from the complexity approach. Thus, it is only at the most complex level that one needs to consider a child's readiness to acquire a new structure.

Note that the sonority sequencing principle also applies to word-final clusters. Here, sonority is expected to fall, as stated in the sonority sequencing principle, and therefore, sonority differences for word-final clusters are negative. Like word-initial clusters, a range of sonority differences are observed in word-final clusters, with smaller differences being more marked than larger differences. It is possible that word-final clusters should be prioritized in a similar manner as word-initial clusters (i.e., prioritize treatment of word-final clusters with a small sonority difference), but to date, there are no studies investigating this claim. Consequently, treatment of clusters should focus on word-initial clusters. Moreover, word-initial clusters are argued to be more complex than word-final clusters (Kirk & Demuth, 2005; Levelt, Schiller, & Levelt, 2000; Lleó & Prinz, 1996).

Characteristics of Children's Knowledge: Accuracy

Gierut, Elbert, and Dinnsen (1987) investigated how knowledge influenced treatment for six children (age = 3;7–4;6) with phonological disorders. Gierut et al. used detailed phonetic and phonemic analyses to categorize each child's knowledge of each singleton consonant, placing each sound on a knowledge continuum from least to most knowledge. Children were then taught three singleton targets sequentially (i.e., Target 1, Target 2, and Target 3) that were either selected from the most to least knowledge end of the continuum (e.g., Target 1-more, Target 2-some, Target 3-least) or the least to most knowledge end of the continuum (e.g., Target 1-least, Target 2-some, Target 3-more). The results showed that children experienced broader system-wide change when treatment began with least knowledge targets. It is noteworthy that the previously described study by Rvachew and Nowak (2001) used a similar definition of knowledge. Recall that their study manipulated knowledge and age-of-acquisition in tandem such that children were taught least-knowledge/late-acquired targets or more-knowledge/early-acquired targets. As noted previously, children taught more-knowledge/early-acquired targets progressed more quickly through the treatment hierarchy and tended to end treatment at sentence production. By comparison, children taught least-knowledge/late-acquired targets progressed more slowly through the treatment hierarchy and tended to end treatment at the word production level. Overall learning appeared equivalent across groups despite having progressed differently through the treatment hierarchy. This finding, in combination with that of Gierut and colleagues (1987), suggests that there is a tension between amount of change and speed of change. That is, quicker progression may be observed for noncomplex targets (more knowledge, early acquired), but broader system-wide change may be observed for complex targets (least knowledge, late acquired). Consistent with the complexity approach, I encourage a focus on broad system-wide change because it would be more likely to result in faster speech sound normalization.

Practically, it is difficult in most common clinical settings to conduct the highly detailed and time-consuming phonological analyses completed by Gierut and colleagues (1987) and Rvachew and Nowak (2001). Accuracy may be an appropriate proxy measure for phonological knowledge and one that is more feasible to collect in everyday clinical settings. In fact, Gierut et al.'s (1987) original continuum does reference concepts that relate to accuracy. For example, at the lowest level of knowledge, the child produces the target incorrectly in all contexts. This would constitute 0% accuracy. Likewise, at the highest level of knowledge, the child produces the target correctly in all contexts. This would constitute 100% accuracy, which would be a target that would not be considered for treatment. The middle levels of knowledge differentiate (from the least to most knowledge) a target that is produced correctly in a few words in a specific word position (e.g., target produced correctly in initial position for some but not all words), a target produced correctly in all words in a specific word position (e.g., target produced correctly in initial position for all words), a target produced correctly in a few words in all word positions (e.g., target produced correctly in initial, medial, and final positions for some but not all words), and a target produced correctly in all positions in all words at least some of the time (e.g., some words alternate between correct and incorrect word productions). In this way, least-knowledge targets would be those with low accuracy (i.e., no correct productions or only very few correct productions). More knowledge targets that would be eligible for treatment would be those with a midlevel of accuracy, which could be confined to a specific word position or could be spread across word positions. As an example, the group of children taught least-knowledge targets by Rvachew and Nowak tended to produce, on average, zero to one item correctly on a 15-item probe (accuracy = 0%–7%) of the target, whereas children taught more knowledge targets tended to produce, on average, three to six items correctly on a 15-item probe (accuracy = 20%–40%). Taken together, accuracy may be an appropriate index of phonological knowledge for clinical practice.

Characteristics of Children's Knowledge: Stimulability

Stimulability is a type of dynamic assessment for phonology. Usually, targets that are produced with low accuracy in a static assessment are examined further (see Powell & Miccio, 1996, for a review). Although there are a variety of approaches to stimulabiltiy testing (Powell & Miccio, 1996), in general, the child is given an accurate model to imitate and sounds are targeted in a variety of potentially facilitative contexts. That is, usually, the child is asked to imitate the target in isolation, in multiple word positions (initial, medial, and final), and with various vowels that may help facilitate correct production. For example, a common approach to stimulability testing for /r/ would require the child to produce the following stimuli in imitation:

r, ri, iri, ir, ra, ara, ar, ru, uru, and *ur* (Miccio, 2002). A target is categorized as stimulable if the child accurately imitates the target three or more times (Miccio, 2002), with some variability across studies in the exact number of correct imitation attempts required (Miccio, Elbert, & Forrest, 1999; Powell, Elbert, & Dinnsen, 1991). If the child only accurately imitates the target fewer times than required (e.g., zero to two), then the target is categorized as nonstimulable. Powell and colleagues (1991) taught six children (age = 4;11–5;6) with phonological disorders /r/ and one other target. The stimulability of all targets of the phonetic inventory was classified (i.e., one or more correct productions on a stimulability task = stimulable, 0 correct productions = nonstimulable). Powell and colleagues observed that, if a child was taught a stimulable target, he or she tended to learn that target and its cognate. In contrast, if a child was taught a nonstimulable target, he or she tended to learn that target and other stimulable sounds. In general, Powell and colleagues concluded that stimulable targets are more likely to be learned on their own without treatment or regardless of the treatment target, whereas nonstimulable targets are unlikely to become accurately produced in the absence of treatment, a conclusion echoed by Miccio and colleagues (1999). Thus, treatment of nonstimulable targets is prioritized within the complexity approach.

Summary: Complex Targets

Taken together, when applying the complexity approach to phonological treatment of 3- to 7-year-old children with phonological disorders, a clinician should prioritize selection of (a) late-acquired, (b) implicationally marked (see Table 1), (c) least-knowledge, and (d) nonstimulable targets to produce broad, system-wide change in phonology.

Implementation: Likely Barriers and Potential Solutions

Even with an understanding of the tenets of the complexity approach, there are at least two likely barriers to implementation. A first potential barrier is that production accuracy and stimulability need to be obtained for each child, and this may or may not be a part of each clinician's standard assessment battery. A second likely barrier is the need to apply and integrate the four pieces of information corresponding to the tenets of the complexity approach: age-of-acquisition, implicational universals, production accuracy, and stimulability. This is a lot of information to keep track of for a potentially large number of targets, especially when singletons and clusters are both considered viable options for treatment. These tasks must be accomplished by a clinician in the context of a potentially large existing workload, where minimal protected time is available for comprehensive assessment and intervention planning.

In the hopes of making implementation of a complexity approach more viable, supplemental materials are provided at KU ScholarWorks, which is the digital repository of the University of Kansas. Materials relevant to this article are available at http://hdl.handle.net/1808/24767 and are licensed under a Creative Commons Attribution-NonCommercial 4.0 International License. This license allows users to share and adapt our materials, on the condition that appropriate credit is given and that the use of the materials is not for commercial purposes. The Creative Commons license allows individual clinicians to tailor the materials to better fit their needs and then share altered materials with their colleagues. The ScholarWorks supplement consists of probes for singleton and cluster production accuracy and stimulability, which are both integrated with a target selection worksheet that quickly and efficiently summarizes the relevant data for treatment target selection. The supplemental materials consist of one Excel workbook with multiple worksheets (1. All_ScoreWorkSheets), two PowerPoint files (2. Cluster_Probe_Pictures and 2. Singleton_Probe_Pictures), and three cases, each with a completed Excel worksheet and video demonstrating completion of the worksheet (3. Case1_4yGirl, 4. Case2_4yBoy, and 5. Case3_4yBoy). Each relevant item is reviewed, in turn, but note that the 1. All_ScoreWorkSheets Excel file contains the following worksheets: (a) "ReadMe": a quick summary of the construction of the materials provided; (b) "Klatt": a description of the computer readable transcription system used in the worksheets; and (c) "Instructions": step-by-step directions for using each resource. Table 3 provides a summary of the time involved in administering each probe and in entering data in the provided Excel sheets.

Singleton Probe

Most clinicians probably use a broad test of articulation as part of their assessment battery for children with phonological disorders. A broad test samples a wide range of targets in the native language but typically in a small number of words or contexts. For example, a broad test may sample each singleton consonant in each word position that it legally occurs in English but only in one word. Thus, for most consonant singletons, you would have one production each in initial, medial, and final positions. This would not be a sufficient sample for computing accuracy because there are not enough opportunities to examine consistency of production in each word position or even overall across word positions. For this reason, a deep articulation test is needed. A deep articulation test samples a target multiple times in each word position so that a clinician can examine consistency of production in each word position and overall. Although there are existing deep tests of articulation, many of these contain many items, requiring too much time for administration, or are available only as word lists, leaving clinicians to find pictures for each item. The included singleton probe targets only the mid- and late-acquired singleton targets /k g f v θ ð s z ʃ ʧ ʤ ŋ l r/, excluding the less frequently occurring /ʒ/ and the less frequently in error /t/. In addition, three singleton targets

Table 3. Estimated administration time for the provided probes and data entry and scoring time for the provided Excel worksheets.

Item	Estimated average time (min)	Estimated range of time (min)
Singleton production probe	25	20–30
Singleton stimulability probe	5	1–10
Cluster production probe	13	10–15
Cluster stimulability probe	5	1–10
Singleton entry and scoring	6	5–10
Cluster entry and scoring	6	5–10
Target selection and generalization prediction entry	13	5–20[a]

[a]Longer times (near 20 min) will likely be needed when using this worksheet for the first time. Shorter times (near 5 min) are more likely when the clinician is more familiar with the steps for completing the worksheet.

relevant to selecting three-element onset clusters also are sampled /p t w/ in a limited manner. Targets are sampled in five words in initial position (with the exception of /ŋ/, which does not occur in initial position) and five words in final position (with the exception of /ð/, which occurs rarely in this position). The three targets relevant to three-element onset clusters are only sampled in five words in initial position, matching the location of the three-element clusters. Last, words that contained targets of interest in initial and final positions were prioritized when creating the probe to increase the efficiency of the probe, although this does lead to selecting a few words that may not be familiar to some children, depending on age and language skills. These procedures yield an 87-word singleton probe that can be administered in approximately 20–30 min (see Table 3).

The Excel workbook (1. All_ScoreWorkSheets) contains the Probe_Score_C worksheet. This worksheet contains a randomized list of the 87 words. The PowerPoint file labeled “2. Singleton_Probe_Pictures” contains the 87 corresponding pictures, which are licensed under Creative Commons and were found at pixabay.com. The spreadsheet shows a prompt to elicit the child’s production of the target word in response to the provided picture. The worksheet also lists which sounds in the word are targets, provides a space for entering a transcription of the child’s production of each target, and provides a space for scoring that production as correct (score = 1) or incorrect (score = 0). I recommend scoring distortions as correct because the complexity approach prioritizes treatment of substitutions or deletions, rather than distortions. Once the probe has been administered and scored, the results are automatically summarized by Excel formulas at the bottom of the spreadsheet. These formulas calculate accuracy in the onset position, in the coda position, and overall across the two positions, which will be used for treatment target selection. In addition, the transcription of the child’s production is listed next to the accuracy calculations so that a clinician can examine error patterns and note these, or any other relevant observations, in Column D. This also allows the clinician to identify distortions and then decide whether to compute accuracy separately for these targets. That is, the clinician may wish to report the accuracy for a target when distortions are counted as correct (as I directed) and/or when distortions are counted as incorrect (the alternative calculation performed by the clinician). Although treatment of distorted targets is not prioritized within the complexity approach, clinicians may want to address these errors for certain children. The clinician could choose to target distortions instead of following a complexity approach or could do so in addition to the complexity approach (e.g., have a goal related to distortions and a goal aligned with the complexity approach). Note that clinicians may wish to copy all or a portion of this summary table into reports to efficiently convey the results of this deep test of articulation. However, I discourage clinicians from deleting rows or columns from any of the Excel worksheets because this may interfere with the correct calculation of other formulas within the Excel workbook. Thus, clinicians will want to manipulate the summary table in a different document, including changing the symbols to orthography to be more parent-friendly.

Singleton Stimulability

The Excel file (1. All_ScoreWorkSheets) contains the Stim_Score_C worksheet. Probe accuracy is automatically copied over from the Probe_Score_C worksheet. Typically, stimulability testing is only performed for targets that are absent from the child’s phonetic inventory (Miccio, 2002). Because our procedures do not involve constructing a phonetic inventory, I recommend testing stimulability for targets with very low accuracy, operationalized as 0%–10% accuracy (i.e., zero or one correct production). To track this, clinicians may want to apply gray shading to the rows for targets with higher accuracy to remind them to not test stimulability for these targets. It also is important to delete the numbers in the summary columns (M, N, O, P) and enter N/A in the final column (P) for untested targets to avoid confusion when using other worksheets that rely on the stimulability data. To test stimulability, I follow the procedures of Miccio (2002), who elicited targets via imitation in multiple word positions and with various preceding or following vowels. For example, as shown in Stim_Score_C, a child would be required to imitate /r/ in isolation and /ri iri ir ra ara ar ru uru ur/. The clinician may choose to modify this procedure and provide greater articulatory cuing, as described by Glaspey and Stoel-Gammon (2005). With either elicitation procedure, the clinician would score each attempt as correct (score of 1) or incorrect (score of 0). The worksheet then automatically summarizes the number correct and percent correct. As described by Miccio, the worksheet codes target as stimulable (score of 1) if 30% or more of the items eliciting that target are imitated correctly and codes target as nonstimulable (score of 0) if fewer than 30% of the items are imitated correctly. Time to administer the stimulability probe varies from 1 to 10 min (see Table 3), depending on the number of low accuracy sounds that need to be tested. Time to enter and score the singleton production probe and the stimulability probe in the Excel worksheet varies from 5 to 10 min (see Table 3).

Cluster Probe

The Excel workbook (1. All_ScoreWorkSheets) contains the Probe_Score_CC(C) worksheet, which is a deep test of two- and three-element onset clusters. Each cluster is targeted in two words. Although this is a small sample of each individual cluster, the analysis focuses more on patterns across classes of clusters: either classes organized by sonority difference, which is important for the complexity approach, or classes organized by a common sound in the cluster (e.g., *l*-clusters, *r*-clusters), which may be useful when sharing results with parents or teachers. As summarized in Table 4, each class is sampled in 4–18 words depending on the specific clusters that fall into a particular class. This yields a 56-word cluster probe that can be administered in approximately 10–15 min (see Table 3). The Probe_Score_CC(C) worksheet is in a format similar to the singleton production probe. The worksheet contains a randomized list of the 56 words, and the target cluster, sonority difference class, and common sound class are listed. The corresponding PowerPoint file 2. Cluster_Probe_Pictures contains the matching 56 pictures. The Probe_Score_CC(C) worksheet shows a prompt for eliciting production of the target word, a space for transcribing the child's production of the target cluster, and a space for scoring the accuracy of the child's production (0 = *incorrect,* 1 = *correct*). As with the singleton probe, I again recommend scoring distortions as correct for the previously stated reasons. Similar to the singleton probe, the results of the cluster probe are automatically summarized by Excel formulas at the bottom of the worksheet. These formulas calculate accuracy by sonority difference class and accuracy by common sound class and list the child's actual productions for each class. Note that the target clusters for each broader class of clusters are listed, and the child's actual productions for those specific clusters are shown in order. For example, for a sonority difference of 6, /kw tw/ are listed as the targeted clusters. The child's production of the two /kw/ items is listed first, followed by the child's production of the two /tw/ items. Thus, the clinician can use this more detailed information to determine which specific clusters are being produced accurately or inaccurately and also can identify distorted productions.

Cluster Stimulability

I could not find any publications where stimulability of clusters was tested. However, a stimulability probe for clusters is provided, but clinicians should be cautioned that there is no specific research about testing stimulability of clusters or applying cluster stimulability to select cluster treatment targets. Thus, it is up to individual clinicians to decide whether stimulability of clusters is useful in their clinical practice. The Excel file (1. All_ScoreWorkSheets) contains the Stim_Score_CC(C) worksheet. Because there are no existing stimulability tests for onset clusters, I follow similar principles to Miccio's (2002) test for singletons in that each target is elicited in isolation and in several nonsense syllable contexts. In some cases, it may be difficult to elicit a cluster in complete isolation. In these cases, use of the mid central vowel /ʌ/ is appropriate (e.g., /tw/ elicited as /twʌ/). Because the position for the cluster is set, namely, onset, multiple word positions do not need to be tested. Instead, a wider array of vowel contexts (i.e., front high, front mid, front low, back high, back mid, back low) are used to facilitate correct production. Probe accuracy is automatically copied over from the Probe_Score_CC(C) worksheet. As with singletons, stimulability would only be tested for clusters with low accuracy. As with the singleton probe, clinicians would shade items not to be tested in gray, remove formulas (in Columns M, N, O, and P), and enter N/A in the final column (P) for clarity in other worksheets that draw data from the cluster stimulability worksheet. I define stimulable and nonstimulable targets with cutoffs similar to Miccio, who defined stimulable targets as those with 3+ correct productions of 10 attempts (30% or greater) and nonstimulable targets as those with 0–2 correct productions of 10 attempts (0%–20%). Thus, individual clusters, which are tested in seven items, are scored as stimulable if two to seven productions are correct (29%–100%) or nonstimulable if zero to one production is correct (0%–14%). As with singleton stimulability, time to administer the cluster stimulability probe varies from 1 to 10 min (see Table 3), depending on the number of low-accuracy clusters that need to be tested. Time to enter and score the cluster production probe and the cluster stimulability probe in Excel varies from 5 to 10 min (see Table 3).

Target Selection

The Excel file (1. All_ScoreWorkSheets) contains the Target_Selection worksheet, which is used to integrate the obtained information about the child's accuracy and stimulability for each target along with information about each target's developmental norms and implicational universals. The first row lists all English singletons and clusters (with the exception of /ʒ/ because of its rare occurrence). Targets that are not elicited in any of our provided probes or are elicited in a limited manner (i.e., /p t w/) are

Table 4. Characteristics of the cluster probe available in the University of Kansas ScholarWorks supplement.

Cluster class	Onset clusters in class	Total words sampled
6	/tw/, /kw/	4
5	/pl/, /kl/, /pr/, /tr/, /kr/	10
4	/bl/, /gl/, /br/, /dr/, /gr/, /sw/	12
3	/fl/, /sl/, /fr/, /θr/, /ʃr/	10
2	/sm/, /sn/	4
−2	/sp/, /st/, /sk/	6
w-clusters	/kw/, /tw/, /sw/	6
l-clusters	/kl/, /pl/, /bl/, /gl/, /fl/, /sl/	12
r-clusters	/kr/, /pr/, /tr/, /br/, /dr/, /gr/, /fr/, /θr/, /ʃr/	18
s-clusters	/sw/, /sl/, /sm/, /sn/, /sp/, /st/, /sk/	14
Three-element clusters	/skw/, /spl/, /skr/, /spr/, /str/	10

shaded in gray, and much of the information for these targets is listed as N/A. I include these items in case clinicians would want to supplement what I have provided and test these early-acquired targets. Sampled targets are shaded in green, with dark green for singletons and light green for clusters.

Developmental Norms

The first two rows of the worksheet show the recommended age-of-acquisition in months for girls (Row 1) and boys (Row 2) based on the norms of Smit et al. (1990). Note that Smit and colleagues did not provide an age-of-acquisition for the cluster /ʃr/. To avoid missing data, I averaged the age-of-acquisition for the other clusters with a sonority difference of 3 and used the result as the age-of-acquisition for /ʃr/. The child's chronological age in months needs to be entered in Cell B18, which is highlighted in yellow. Once the age is entered, formulas in Rows 4 and 5 compute the difference between the child's age and the recommended age-of-acquisition for each target (i.e., recommended age-of-acquisition − child's chronological age). A negative number means that the recommended age-of-acquisition is younger than the child's chronological age. A positive number means that the recommended age-of-acquisition is older than the child's chronological age. Rows 6 and 7, highlighted in blue, use this difference to score each target as late-acquired (score of 1) or not (score of 0). A target is scored as 0 if the difference score is at or below +12 months, capturing targets that should have been acquired by the child's current age (negative scores to 0) or that should be acquired within the next year (scores of 0 to +12). In contrast, a target will be scored as 1 if the difference score is above +12 months, capturing targets that a child would not be expected to acquire in the near future. Note that this coding matches the definition of late acquired used by Gierut and colleagues (1996), who defined late-acquired targets as those having a score of +12 or greater. For ease of reading, clinicians may want to clear the information for the gender that is not relevant for their target child. For example, for a female child, the information relevant to male children (Rows 3, 5, and 7) would be cleared, meaning that the rows remain but the information is deleted. The provided case videos demonstrate this.

Implicational Universals

For implicational universals, everything is provided in the worksheet and no further data entry is needed from the clinician. Row 8 shows the number of different implicational universals that are relevant to a given target. For example, for /f/, there is one relevant implicational universal: Fricatives imply stops. Thus, the fricative /f/ is more marked than stops and receives a code of 1. On the other hand, there are three relevant implicational universals for /ʤ/: (a) Affricates imply fricatives, (b) fricatives imply stops, and (c) voiced obstruents imply voiceless obstruents. Thus, the code for /ʤ/ is 3. Supplemental Material S1 in the journal supplemental materials shows the details of how each target was coded. Row 9 then provides the implicational universal score. Implicational universals are scored separately for singletons and clusters because most studies have examined treatment of singletons and clusters separately. The implicational universal score was determined using a median split of the codes for singletons and clusters separately. Specifically, the codes for singletons were 0, 1, 2, and 3. Thus, singletons with a code of 0 or 1 were scored as 0 = *less complex,* and those with a code of 2 or 3 were scored as 1 = *more complex*. The codes for clusters were 0, 3, 4, 5, 6, 7, and 8. Clusters with a code of 0, 3, or 4 were coded as 0 = *less complex,* and those with a code of 5, 6, 7, or 8 were coded as 1 = *more complex*. This also matches how Gierut coded cluster complexity in her treatment study (Gierut, 1999). In addition, on the basis of Gierut's observation that children may group /s/ + nasal clusters with adjuncts because of their superficial similarity (i.e., both are *s*-clusters), I marked /sm sn/ as N/A in multiple places on the worksheet because more research is needed before these clusters are selected for treatment.

Accuracy

Row 10 copies the accuracy from the singleton and cluster probe summaries. Thus, accuracy will automatically appear here, and no data entry is required. Likewise, Row

11 scores the accuracy as 0 = *higher accuracy* for targets with an overall accuracy of 11% or higher and 1 = *low accuracy* for targets with an overall accuracy of 0%–10%.

Stimulability

Like accuracy, stimulability is automatically entered by formula, and no data entry is required. Specifically, Row 12 copies the stimulability code, 0 = *nonstimulable* (zero to two correct singleton productions, zero to one correct cluster productions), 1 = *stimulable* (three or more correct singleton productions, two or more correct cluster productions), or N/A = *not tested,* from the stimulability worksheets. Row 13 then automatically scores this as 1 = *nonstimulable* or 0 = *stimulable or N/A*.

Total Score

Finally, Row 14 sums the developmental norms (Row 6 or 7 depending on gender), implicational universals (Row 9), accuracy (Row 11), and stimulability (Row 13) scores. Thus, a target that is late acquired, marked based on a high number of relevant implicational universals, of low accuracy, and nonstimulable would receive a total score of 4, which is the highest score possible. These targets with a score of 4 represent the most complex targets that could be selected for treatment. In complement, a target that is early acquired, unmarked based on a low number of relevant implicational universals, of high accuracy, and N/A for stimulability (because stimulability was not tested because of high accuracy) would receive a total score of 0, which is the lowest score possible. These targets would not be selected for treatment based on the complexity approach. Targets receiving a total score of 1–3 obviously fall in the middle of these two extremes and potentially could be selected for treatment within the complexity approach because they are complex on some dimensions. Moreover, clinicians consider a range of factors (Powell, 1991) when selecting targets for treatment (e.g., parent and child goals), and those factors would need to be integrated with the information about complexity. That is, there may be reasons beyond the complexity approach to select a target that has a score less than 4, and this would be an appropriate way to integrate complexity with other factors (e.g., needs of the child). Cases will be used to demonstrate the decision-making process, but there are two caveats related to clusters that warrant comment.

The first caveat related to clusters is only a reminder that it is up to the clinician to decide whether to test stimulability of clusters because there is minimal guidance in the literature. If a clinician does not test stimulability of clusters, then the maximum total score on the worksheet for clusters is 3, rather than 4. A clinician would only need to keep this in mind when comparing total scores of clusters and singletons and may even want to change the scores of 4 for singletons and scores of 3 for clusters to “max” to make this transparent and avoid confusion. The second caveat relates to the selection of three-element clusters. Recall that the study by Gierut and Champion (2001) showed that there was greater learning for children who were taught a three-element cluster where the second and third sounds of the cluster were known by the child. It was argued that knowledge of the second and third sounds focused attention on the true cluster element promoting greater system-wide change. In complement, there was less learning for children who were taught a three-element cluster where the first and second sounds of the cluster were known by the child, presumably because this focused attention on the adjunct element of the cluster. Thus, it is important to consider children’s knowledge of each element of a three-element cluster before selecting a three-element cluster as a treatment target. The target selection worksheet summarizes singleton accuracy in onset position for each sound in each three-element cluster (Columns AX–BD) in Rows 19–21. Note that Gierut and Champion’s measure of knowledge was a phonemic inventory. It is not straightforward to translate this definition of knowledge into one based on accuracy. Thus, the worksheet notes that higher accuracy of the second and third elements in the three-element cluster is desirable but does not provide a specific cut-point. Personally, I would select 40% (two correct productions in onset position) or higher because this would suggest at least emerging knowledge of the sound as a singleton without requiring mastery. Whatever cutoff is chosen, if the child shows higher accuracy for /s/, which is always the first element in a three-element cluster, then all three-element clusters should be eliminated from consideration as treatment targets by changing the cluster’s total score in Row 14 from a number to “N/A.” If three-element clusters are still a viable option, then the next step is to identify three-element clusters that have a low accuracy for the second and third elements in the cluster and eliminate these from consideration as treatment targets. It is possible that no three-element clusters will remain as potential treatment targets.

Predicting Generalization

One last resource is provided to assist clinicians in selecting between multiple potential treatment targets. The Excel file (1. All_ScoreWorkSheets) contains the GeneralizationPrediction worksheet. This worksheet is adapted from Gierut and Hulse (2010), who provided a hard copy worksheet for identifying which low-accuracy singletons would likely change if different singletons were selected for treatment. The generalization prediction was based on implicational universals. For example, if the voiced affricate /ʤ/ was selected for treatment, then low-accuracy voiceless affricate, voiced and voiceless fricatives, and voiced and voiceless stops would be predicted to potentially improve in accuracy. I build on the work of Gierut and Hulse (2010) by implementing their singleton worksheet in an electronic format and adding in the clusters, both as potential treatment targets and as sounds that may improve in accuracy through broad system-wide phonological change.

In the GeneralizationPrediction worksheet, potential treatment targets appear in the columns if their total target selection score (Row 14) is 3 or 4. Singletons and clusters that the child produced with 50% or lower accuracy are shown in the rows. The matrix scores a 1 for any lower-accuracy singleton or cluster (i.e., the rows) that is predicted to improve based on implicational universals, if a particular treatment target (i.e., the columns) is selected. The bottom row then totals the number of lower-accuracy singletons and clusters that may improve for each potential treatment target. This total provides a way to compare the predicted impact of different potential treatment targets. It is important to note that this is only a prediction, and it is only based on implicational universals. Thus, it is important that clinicians collect data to verify what the child is actually learning as treatment progresses and at the end of treatment. Given the importance of data collection, it is critical that the words used on the provided probes not be selected for treatment so that the probe will remain a valid assessment of generalization beyond the treatment targets. Time to enter data in the target selection and generalization worksheets varies from 5 to 20 min (see Table 3), depending on familiarity with the steps involved.

Case Illustrations

Three cases are provided in the ScholarWorks supplement. Each case includes an Excel file, which is the completed 1. All_ScoreWorkSheets, and a video showing how the 1. All_ScoreWorkSheets was completed. The cases are based on three children who were seen as part of a research study. Child 1 has a complete phonological battery that includes the singleton and cluster probes as well as stimulability for both singletons and clusters. Child 2 includes the singleton and cluster probes but not stimulability testing because the child did not qualify for the research study. Consequently, the full battery was not administered. Child 3 was seen before the development of these materials but was administered a comprehensive singleton probe as well as stimulability for singletons. Thus, his partial data focusing on singletons could be used to illustrate selection of singleton treatment targets. Children 2 and 3 demonstrate that clinicians can choose which components of the provided materials they use. That is, it is not required that all provided materials be administered to every child. Table 3 may be useful in weighing the cost (in time) of using each item.

Child 1

Child 1 is a 4-year-, 4-month-old girl who scored at the sixth percentile on the Goldman-Fristoe Test of Articulation–Third Edition (Goldman & Fristoe, 2015). Figure 1 shows the summary from the singleton probes. Child 1 showed low accuracy for /θ ð r/ and was nonstimulable for these targets. Figure 2 shows the summary of accuracy for the cluster probe. Here, the child had difficulty with two-element *r*-clusters, only producing one target (i.e., /kr/) accurately one time. Child 1 also had consistent difficulty with three-element *r*-clusters. In addition to *r*-clusters, Child 1 had difficulty producing clusters containing /t/ (i.e., /tw tr str/), where she substituted [k] for target /t/. This [k] for /t/ substitution was not observed in singletons. Child 1 was nonstimulable for all erred clusters, except /tw/.

Figure 3 shows a portion of her target selection worksheet (top) and predicted generalization worksheet (bottom). Among the singletons, the highest total complexity score (top of Figure 3) was 3, which was obtained for /θ r/. /θ r/ were late acquired, of low accuracy, and nonstimulable but not marked based on implicational universals. Among the singletons, these are the most complex treatment target options for Child 1. In terms of clusters, the highest total complexity score was 4, which was obtained for /br dr gr fr θr ʃr/. These potential treatment targets were late acquired, of low accuracy, marked based on implicational universals, and nonstimulable. Note that three-element clusters were not a viable treatment option for this child, primarily due to her high accuracy in producing /s/, which could focus her on the adjunct portion of the cluster.

A clinician might choose to select one of the cluster targets over the singleton targets because the clusters received a total complexity score that was higher than the singleton targets. The predicted generalization, shown in the bottom of Figure 3, further supports the potential impact of selecting a cluster for treatment, rather than a singleton. Specifically, if /fr/, /θr/, or /ʃr/ were selected for treatment, almost all of Child 1's errors are predicted to improve. The only erred sound not predicted to improve is /r/. There is evidence that children may acquire a distinction between /l/ and /r/ when they acquire liquid clusters, but this is based on somewhat limited evidence (Gierut & O'Connor, 2002). Specifically, in a retrospective study, Gierut and O'Connor (2002) observed that 100 children had no liquid distinction and no liquid clusters, three children had a liquid distinction and no liquid clusters, and two children had both a liquid distinction and liquid clusters. These patterns are consistent with the hypothesis that a liquid distinction is less marked relative to liquid clusters. However, five children showed a pattern counter to the hypothesis: production of liquid clusters but no liquid distinction. Thus, a clinician might tentatively hypothesize that selection of an *r*-cluster could improve production of singleton /r/, but this hypothesis would need to be closely monitored.

Taken together, treatment of /fr/, /θr/, or /ʃr/ is the most complex treatment option for this child and is predicted to produce the greatest phonological change. The clinician would now appeal to other factors to select one specific cluster of these three options. In addition, clusters /br dr gr/ only differ slightly in predicted generalization from /fr θr ʃr/. Consequently, if there were a compelling reason outside the complexity approach to select one of these targets, then selection of /br/, /dr/, or /gr/ might be appropriate. For Child 1, I would select only one of these complex targets for two reasons. One reason is that selecting only one target allows for higher treatment intensity

Figure 1. Singleton accuracy (top) and singleton stimulability (stim, bottom) for Child 1.

Accuracy Analysis	Onset #	Onset %	Coda #	Coda %	Total %	Onset Productions					Coda Productions				
k	5	100%	5	100%	100%	k	k	k	k	k	k	k	k	k	k
g	5	100%	5	100%	100%	g	g	g	g	g	g	g	g	g	g
f	5	100%	5	100%	100%	f	f	f	f	f	f	f	f	f	f
v	5	100%	5	100%	100%	v	v	v	v	v	v	v	v	v	v
T (θ)	0	0%	1	20%	10%	p	d	st	f	f	f	f	f	T	f
D (ð)	1	20%	1	33%	25%	D	d	d	d	d	D	v	v	N/A	N/A
s	5	100%	4	80%	90%	s	s	s	s	s	S	s	s	s	s
z	5	100%	4	80%	90%	z	z	z	z	z	z	z	Z	z	z
S (ʃ)	5	100%	5	100%	100%	S	S	S	S	S	S	S	S	S	S
C (ʧ)	5	100%	5	100%	100%	C	C	C	C	C	C	C	C	C	C
J (ʤ)	4	80%	5	100%	90%	g	J	J	J	J	J	J	J	J	J
G (ŋ)	N/A	N/A	3	60%	60%	N/A	N/A	N/A	N/A	N/A	n	n	G	G	G
l	7	100%	5	100%	100%	l	l	l	l	l	l	l	l	l	l
r	0	0%	0	0%	0%	w	w	w	w	w	del	del	del	del	del
ALL	**52**	**78%**	**53**	**78%**	**78%**	67 opportunities					68 opportunities				

Probe Accuracy	Target Sound	Isolation	#_i	i_i	i_#	#_a	a_a	a_#	#_u	u_u	u_#	# Correct	Total	%	Stim?
100%	f														N/A
100%	v														N/A
10%	T (θ)	1	0	0	0	0	0	0	0	0	0	1	10	10%	0
25%	D (ð)	1	0	0	0	0	0	0	0	0	0	1	10	10%	0
100%	C (ʧ)														N/A
90%	J (ʤ)														N/A
60%	G (ŋ)														N/A
100%	l														N/A
0%	r	0	0	0	0	0	0	0	0	0	0	0	10	0%	0

than selecting multiple targets. That is, the entire session can focus on the single target, increasing the number of production trials and feedback devoted to that target. A second reason is that these target options predict generalization to the same sounds. Thus, selecting multiple targets from this set would be redundant and likely unnecessary. It is important to note that clinicians may be challenged to justify their selection of a complex target, like /fr θr ʃr/, for a preschool child. Clinicians can respond to such challenges by noting the relevant research evidence, in this case Gierut (1999), showing that treatment of complex targets, even in preschool children (aged 3;8–7;8 in Gierut, 1999), results in broad, system-wide change in phonology. In addition, the clinician should note that she or he will be providing support and coaching to minimize the child's frustration and maximize the child's success in producing the complex target. There are numerous resources for clinicians to gain insights into how to support correct production of many targets (e.g., Bauman-Wängler, 2012, see chapter on phonetic placement; Bleile, 2018; Secord, Boyce, Donohue, Fox, & Shine, 2007; SLPath.com).

Child 2

Child 2 is a 4-year-, 4-month-old boy who scored at the < 0.1 percentile on the Goldman-Fristoe Test of Articulation–Third Edition (Goldman & Fristoe, 2015). His full analysis is available in the ScholarWorks supplement (4. Case2_4yBoy). Child 2 showed very low accuracy for singletons /k g f θ ð ʤ l r/, emerging accuracy for singletons /v ʧ ŋ/, and very low accuracy for all clusters except /st/. Recall that stimulability testing was not performed for this child. Thus, the highest possible complexity score

Figure 2. Cluster accuracy for Child 1. CC = consonant–consonant; CCC = consonant–consonant–consonant; SD = sonority difference.

SD	#	%	Accuracy Analysis by SD	Productions											
6	2	50%	kw, tw	kw	kw	kw	t								
5	5	50%	kl, pl, kr, pr, tr	kl	kl	pl	pl	kr	kw	pw	pw	kw	kw		
4	6	50%	bl, gl, br, dr, gr, sw	bl	bl	gl	gl	bw	bw	gw	gw	gw	gw	sw	sw
3	4	40%	fl, sl, fr, Sr (ʃr), Tr (θr)	fl	fl	sl	sl	f	fw	Sw	Sw	f	fw		
2	4	100%	sm, sn	sm	sm	sn	sn								
-2	6	100%	sk, sp, st	sk	sk	sp	sp	st	st						
3s	4	40%	skw, spl, skr, spr, str	skw	skw	spl	spl	skw	skw	spw	sp	sk	skw		

Sound	#	%	Accuracy Analysis by Sound	Productions																	
w-cluster	**4**	**67%**	**kw, tw, sw**	**kw**	**kw**	**kw**	**t**	**sw**	**sw**												
l-cluster	12	100%	kl, pl, bl, gl, fl, sl	kl	kl	pl	pl	bl	bl	gl	gl	fl	fl	sl	sl						
r-cluster	**1**	**6%**	**kr, pr, tr, br, dr, gr, fr, Sr (ʃr), Tr (θr)**	**kr**	**kw**	**pw**	**pw**	**kw**	**kw**	**bw**	**bw**	**gw**	**gw**	**gw**	**gw**	**f**	**fw**	**Sw**	**Sw**	**f**	**fw**
s-cluster	14	100%	sw, sl, sm, sn, sk, sp, st	sw	sw	sl	sl	sm	sm	sn	sn	sk	sk	sp	sp	st	st				
3s	**4**	**40%**	**skw, spl, skr, spr, str**	**skw**	**skw**	**spl**	**spl**	**skw**	**skw**	**spw**	**sp**	**sk**	**skw**								

Summary		
True CC	21	53%
Adjunct	6	100%
CCC	4	40%
All	31	55%

is 3 because only developmental norms, implicational universals, and accuracy can be considered in selecting targets. Figure 4 shows a portion of his target selection worksheet (top) and predicted generalization worksheet (bottom). Among the singletons, /ð ʤ/ achieved the highest total complexity score (top of Figure 4) because they were late acquired, marked based on implicational universals, and of low accuracy. In terms of clusters, the highest total complexity score was obtained for /bl gl sw br dr gr fl sl fr θr ʃr/. These potential treatment targets were late acquired, marked based on implicational universals, and of low accuracy. The predicted generalization (bottom of Figure 4) suggests that /ð/ may not be as a good of a choice as the other options because less generalization is predicted. In contrast, treatment of /ʤ/ may produce change in the affricates, fricatives, and stops that are produced in error. Treatment of one of the clusters also is predicted to lead to change in these singletons as well as change in other clusters. Taken together, /ʤ bl gl sw br dr gr fl sl fr θr ʃr/ are viable complex targets for Child 2 that are predicted to spark broad, system-wide change.

In choosing a final target for Child 2, a clinician would definitely want to consider Child 2's behavior. Child 2 would sit and attend well for only short periods, and his behavior declined across the session. A clinician would likely see this child for sessions no longer than 20 min. A clinician would also want to consider ease of teaching the target because Child 2 would likely not be attentive to extensive feedback or coaching, and that could further limit the intensity of the session by reducing the number of trials achieved in 20 min (or less). A clinician might consider selecting /bl/ as the treatment target because Child 2 produces /b/ accurately and produces [bw] as a substitute for other cluster targets (including target /bl/), indicating that the child has an emerging ability to combine /b/ with a sound to create a cluster. In contrast, Child 2 typically reduces clusters to singletons, with [bw] being the only example of cluster simplification for a true cluster target (but see the adjunct clusters). It is important to note that this level of cluster reduction may not be developmentally appropriate. Smit (1993b, p. 945) notes that "reduction of many clusters is no longer typical by the age of 3;6, and that remaining clusters are preserved as clusters by age 4;0–5;0, although there may continue to be segmental errors within these clusters" (see also McLeod, van Doorn, & Reed, 2001, for converging data). Note that this information about typical error patterns for clusters can be used to justify selection of /bl/, if challenged. In terms of further justification for /bl/ as a treatment target for Child 2, /l/ is visible and Child 2 seemed to respond to visual cuing for other tasks. Likewise, as noted for Child 1, treatment of /bl/ could facilitate learning of /l r/ as singletons. Thus, /bl/ might be a treatment target that would integrate the complexity approach with the child's needs as well as his current skills and abilities. Clinicians likely would have additional hypotheses about which of these complex targets is a good fit for Child 2 (e.g., /fl/ is also visible and occurs in many common words). That is, /bl/ is not the only well-reasoned target for Child 2, but its selection illustrates how a clinician can appeal to multiple factors beyond the

Figure 3. Partial example of target selection worksheet (top) and predicted generalization worksheet (bottom) for Child 1. AoA = age of acquisition; CA = chronological age; CC = consonant–consonant.

	S(ʃ)	T(θ)	s	r-	-r	D(ð)	v	C(ʧ)	z	J(ʤ)	bl	gl	sw	br	dr	gr	fl	sl	fr	T(θ) +r	S(ʃ) +r	skw	spl	spr	str	skr
Female AoA - Child's CA	**20**	**20**	**32**	**44**	**44**	**2**	**14**	**20**	**32**	**20**	**14**	**14**	**32**	**44**	**44**	**44**	**14**	**32**	**44**	**56**	**37**	**32**	**32**	**32**	**32**	**32**
Female Norms Score (1 = late acquired)	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1			1		
Implicational Universals (# of laws that apply)	**1**	**1**	**1**	**1**		**2**	**2**	**2**	**2**	**3**			**5**						**6**					**8**		
Implicational Score (1 = complex)	0	0	0	0		1	1	1	1	1			1						1					1		
Accuracy (Total % Correct)	**100%**	**10%**	**90%**	**0%**		**25%**	**100%**	**100%**	**90%**	**90%**			**50%**						**40%**					**40%**		
Accuracy Score (1 = low accuracy)	0	1	0	1		0	0	0	0	0			0						0					0		
Stimulability (0 = non; 1 = stim)	**N/A**	**0**	**N/A**	**0**		**0**	**N/A**	**N/A**	**N/A**	**N/A**	**N/A**	**N/A**	**N/A**	**0**	**0**	**0**	**N/A**	**N/A**	**0**	**0**	**0**	**N/A**	**N/A**	**0**	**0**	**0**
Stimulability Score (1 = nonstimulable)	0	1	0	1		1	0	0	0	0	0	0	0	1	1	1	0	0	1	1	1	0	0	1	1	1
Total Score (4 = Best, Highest)	**N/A**	**3**	**N/A**	**3**	**3**	**2**	**N/A**	**N/A**	**N/A**	**N/A**	**N/A**	**N/A**	**N/A**	**4**	**4**	**4**	**N/A**	**N/A**	**4**	**4**	**4**	**N/A**	**N/A**	**N/A**	**N/A**	**N/A**
Individual CC Accuracy	**N/A**	**N/A**	**N/A**	**N/A**	**N/A**	**N/A**	**N/A**	**N/A**	**N/A**	**N/A**	**100%**	**100%**	**100%**	**0%**	**0%**	**0%**	**100%**	**100%**	**0%**	**0%**	**0%**	**100%**	**100%**	**0%**	**0%**	**0%**
Individual CC Score (1 = low accuracy)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0	0	0	1	1	1	0	0	1	1	1	0	0	1	1	1

Low-Accuracy Sounds </= 50%	Potential Treatment Targets								
	theta	r-	-r	br	dr	gr	fr	theta+r	esh+r
T (θ)	1			1	1	1	1	1	1
D (ð)				1	1	1	1	1	1
r		1	1						
kw, tw				1	1	1	1	1	1
kl, pl, kr, pr, tr				1	1	1	1	1	1
bl, gl, br, dr, gr, sw				1	1	1	1	1	1
fl, sl, fr, Sr, Tr							1	1	1
skw, spl, skr, spr, str									
Total # of Low-Accuracy Sounds Predicted to Change for Each Potential Treatment Target	**1**	**1**	**1**	**5**	**5**	**5**	**6**	**6**	**6**

complexity approach to tailor the treatment to an individual child.

Child 3

Child 3 is a 4-year-, 9-month-old boy who scored at the second percentile on the Goldman-Fristoe Test of Articulation–Second Edition (Goldman & Fristoe, 2000). His full analysis is available in the ScholarWorks supplement (5. Case3_4yBoy). Child 3 showed very low accuracy for singletons /k g ð ʃ ŋ l r/ and emerging accuracy for singletons /f θ s z ʧ ʤ/. Child 3 was stimulable for correct production of /ʃ ʧ ʤ/ but was not stimulable for correct production of /k g θ ð l r/. Clusters were not tested as part of the battery for the research study Child 3 participated in. Thus, target selection focuses exclusively on singleton options. Figure 5 shows a portion of Child 3's target selection worksheet (top) and predicted generalization worksheet (bottom). In terms of target selection, /ð/ achieved the highest total complexity score (4; see top of Figure 5) because it was late acquired, marked based on implicational universals, of low accuracy, and nonstimulable. An additional option with a total score of 3 was /l/, which was late acquired, of low accuracy, and nonstimulable. All other singletons had a total score of 2 or less. The predicted generalization (bottom of Figure 5) suggests that selecting both targets might be the best option. Treatment of /ð/ is predicted to affect the stops and fricatives but not influence the nasals or liquids. In contrast, treatment of /l/ is hypothesized to promote learning of the liquid /l/ and affect the nasals. In this way, treatment of both targets could lead to improvements in most of the sounds Child 3 is having difficulty with (the exception being the affricates). Although I generally recommend selecting only one treatment target, there can be cases, like Child 3, where it may make sense to select two treatment targets to potentially produce change in as many low-accuracy sounds as possible. Another option is that the clinician may decide to examine cluster production because there were not many complex targets among the singletons and correct production is emerging for a number of the singletons. Likewise, Child 3 showed a pattern of cluster reduction for many of the cluster targets on the Goldman-Fristoe Test of Articulation–Second Edition, suggesting that clusters may warrant attention for this child because of the error pattern not being typical for his age (McLeod et al., 2001; Smit, 1993b).

Another aspect of Child 3's profile that warrants comment is his distortions. Child 3 lateralized /s z ʧ ʤ/ and produced lateralized /s/ as a substitute for /ʃ/. Recall that distortions are counted as correct because the complexity approach focuses on phonological errors, namely,

Figure 4. Partial example of target selection worksheet (top) and predicted generalization worksheet (bottom) for Child 2. AoA = age of acquisition; CA = chronological age; CC = consonant–consonant.

	k	f-	-f	g	l-	-l	T (θ)	r-	-r	D (ð)	tw	kw	J (ʤ)	pl	kl	pr	tr	kr	bl	gl	sw	br	dr	gr	fl	sl	fr	T(θ) +r	S(ʃ) +r
male developmental norms (late = good)	42	42	66	48	72	84	96	96	96	84	66	66	84	72	72	96	96	96	72	72	84	96	96	96	72	84	96	108	90
Male AoA - Child's CA	-10	-10	14	-4	20	32	44	44	44	32	14	14	32	20	20	44	44	44	20	20	32	44	44	44	20	32	44	56	38
Male Norms Score (1 = late acquired)	**0**	**0**	**1**	**0**	**1**	**1**	**1**	**1**	**1**	**1**	**1**		**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
Implicational Universals (# of laws that apply)	0	1		1	1		1	1		2	3		3	4					5						6				
Implicational Score (1 = complex)	**0**	**0**		**0**	**0**		**0**	**0**		**1**	**0**		**1**	**0**					**1**						**1**				
Accuracy (Total % Correct)	10%	10%		0%	0%		0%	0%		0%	0%		0%	0%					0%						0%				
Accuracy Score (1 = low accuracy)	**1**	**1**		**1**	**1**		**1**	**1**		**1**	**1**		**1**	**1**					**1**						**1**				
Total Score (3 = Best, Highest)	**1**	**1**	**2**	**1**	**2**	**2**	**2**	**2**	**2**	**3**	**2**	**2**	**3**	**2**	**2**	**2**	**2**	**2**	**3**	**3**	**3**	**3**	**3**	**3**	**3**	**3**	**3**	**3**	**3**
Individual CC Accuracy	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0%	0%	N/A	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Individual CC Store (1 = low accuracy)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	1	1	N/A	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Low-Accuracy Sounds </= 50%	Potential Treatment Targets												
	D (ð)	J (ʤ)	bl	gl	sw	br	dr	gr	fl	sl	fr	Tr	Sr
k	1	1	1	1	1	1	1	1	1	1	1	1	1
g	1	1	1	1	1	1	1	1	1	1	1	1	1
f	1	1	1	1	1	1	1	1	1	1	1	1	1
v		1	1	1	1	1	1	1	1	1	1	1	1
T (θ)	1	1	1	1	1	1	1	1	1	1	1	1	1
D (ð)	1	1	1	1	1	1	1	1	1	1	1	1	1
C (ʧ)		1	1	1	1	1	1	1	1	1	1	1	1
J (ʤ)		1	1	1	1	1	1	1	1	1	1	1	1
G (ŋ)													
l													
r													
kw, tw			1	1	1	1	1	1	1	1	1	1	1
kl, pl, kr, pr, tr			1	1	1	1	1	1	1	1	1	1	1
bl, gl, br, dr, gr, sw			1	1	1	1	1	1	1	1	1	1	1
fl, sl, fr, Sr, Tr									1	1	1	1	1
sm, sn													
sk, sp, st													
skw, spl, skr, spr, str													
Total # of Low-Accuracy Sounds Predicted to Change for Each Potential Treatment	**5**	**8**	**11**	**11**	**11**	**11**	**11**	**11**	**12**	**12**	**12**	**12**	**12**

substitutions and deletions. Thus, one might hypothesize that Child 3 is learning /s z ʧ ʤ/ as phonemes (in word-final position) but is not producing the correct target phone, indicating an articulatory error. Although we are focusing on phonological patterns within a complexity approach, it is still important to consider distortions and whether they require attention for a given child. Smit (1993a; Figures 2 and 3) provides normative data on /s/ distortions. In general, dentalized /s/ is quite common throughout the ages studied (2;6–9;0), but lateralized /s/ is relatively rare over the same period. In contrast, Shriberg (1993) classifies both lateralized and dentalized sibilant fricatives and affricates as common clinical distortions. He considers the lateralized distortion a concern at the age of 7 years and beyond and considers the dentalized distortion a concern at the age of 9 years and beyond (see Table A in the appendix of the article). Thus, there is controversy about when to address distortions. If a clinician chose to write a treatment goal targeting the lateralized distortions, she or he would want to consider how to integrate that goal with the complexity goal. That is, the clinician would want to determine whether the two goals should be addressed simultaneously or sequentially, keeping in mind the need to achieve appropriate treatment intensity. The previously referenced materials on eliciting correct productions can be helpful in targeting distortions, but clinicians may also want to consult distortion-specific resources (e.g., Marshalla, 2007).

A final point regarding Child 3 is that the predicted generalization (bottom of Figure 5) suggests that several targets not identified as complex during target selection could lead to broad, system-wide change. Specifically, the

Figure 5. Partial example of target selection worksheet (top) and predicted generalization worksheet (bottom) for Child 3. AoA = age of acquisition; CA = chronological age.

	k	G (ŋ)	f-	-f	g	l-	-l	S (ʃ)	T (θ)	s	r-	-r	D (ð)	C (ʧ)	z	J (ʤ)
male developmental norms (late = good)	42	84	42	66	48	72	84	84	96	84	96	96	84	84	84	84
Male AoA - Child's CA	-15	27	-15	9	-9	15	27	27	39	27	39	39	27	27	27	27
Male Norms Score (1 = late acquired)	**0**	**1**	**0**	**0**	**0**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
Implicational Universals (# of laws that apply)	0	0	1		1	1		1	1	1	1		2	2	2	3
Implicational Score (1 = complex)	**0**	**0**	**0**		**0**	**0**		**0**	**0**	**0**	**0**		**1**	**1**	**1**	**1**
Accuracy (Total % Correct)	0%	0%	50%		0%	0%		0%	20%	50%	0%		0%	50%	50%	50%
Accuracy Score (1 = low accuracy)	**1**	**1**	**0**		**1**	**1**		**1**	**0**	**0**	**1**		**1**	**0**	**0**	**0**
Stimulability (0 = non; 1 = stim)	0	N/A	N/A		0	0		1	0	N/A	1		0	1	N/A	1
Stimulability Score (1 = nonstimulable)	**1**	**0**	**0**		**1**	**1**		**0**	**1**	**0**	**0**		**1**	**0**	**0**	**0**
Total Score (4 = Best, Highest)	**2**	**2**	**0**	**0**	**2**	**3**	**3**	**2**	**2**	**1**	**2**	**2**	**4**	**2**	**2**	**2**

Low-Accuracy Sounds </= 10%	Potential Treated Sounds															
	k	eng	f-	-f	g	l-	-l	S (ʃ)	T (θ)	s	r-	-r	D (ð)	C (ʧ)	z	J (ʤ)
k	1		1	1	1			1	1	1			1	1	1	1
g			1	1	1			1	1	1			1	1	1	1
f			1	1									1	1	1	1
T (θ)									1				1	1	1	1
D (ð)													1	1		1
s										1			1	1	1	1
z														1	1	1
S (ʃ)								1					1	1	1	1
C (ʧ)														1		1
J (ʤ)																1
G (ŋ)		1				1	1				1	1				
l						1	1									
r											1	1				
Total # of Low-Accuracy Sounds Predicted to Change for Each Potential Treatment Target	**1**	**1**	**3**	**3**	**2**	**2**	**2**	**3**	**3**	**3**	**2**	**2**	**7**	**9**	**7**	**10**

affricates /ʧ ʤ/ are associated with the largest predicted change. These were not identified as potential complex targets because of the child's knowledge of these targets. That is, the child produced both affricates with 50% accuracy and was stimulable for correct (albeit lateralized) production. These two pieces of information suggest that the child is learning these targets on his own and treatment may not be needed to facilitate continued growth in accuracy of these sounds. Therefore, I would suggest that the affricates not be selected as a treatment target for Child 3 initially but the accuracy should be monitored to determine whether acquisition of these sounds is continuing or stalling and to track potential change in the lateral distortion (if not immediately targeted in treatment).

Conclusion

Despite the strong evidence base for the complexity approach (Baker & McLeod, 2011; Kamhi, 2006), few clinicians seem to implement this approach in their own clinical practice, likely because of lack of familiarity (Brumbaugh & Smit, 2013). Research suggests that treatment of late-acquired, marked (based on implicational universals), low-accuracy, or nonstimulable targets can trigger broad, system-wide change in phonology. Resources were developed to assist clinicians in implementing the complexity approach in their clinical practice by providing relatively quick but comprehensive deep tests of phonology and stimulability and integrating these assessments with target selection and generalization prediction worksheets. Although using these resources will likely add time to the assessment and treatment planning processes (see Table 3), the effectiveness of using the complexity approach may reduce overall time in treatment, making the initial investment worthwhile. Data from three preschool children illustrate how these resources can be used to identify complex targets and to tailor the final selection of a complex target to a child's needs and abilities. I hope that this tutorial will expand the range of evidence-based treatment options that clinicians can use when treating children with phonological disorders.

Acknowledgments

Adrienne Latta assisted with data collection for two of the three cases. Jill Hoover assisted with data collection for one of the three cases.

References

Baker, E., & McLeod, S. (2011). Evidence-based practice for children with speech sound disorders: Part 1 narrative review. *Language, Speech, and Hearing Services in Schools, 42*(2), 102–139. https://doi.org/10.1044/0161-1461(2010/09-0075)

Barlow, J., & Gierut, J. A. (2002). Minimal pair approaches to phonological remediation. *Seminars in Speech and Language, 23*(1), 54–68.

Bauman-Wängler, J. A. (2012). *Articulatory and phonological impairments: A clinical focus* (4th ed.). Boston, MA: Pearson.

Bleile, K. M. (2018). *The late eight* (3rd ed.). San Diego, CA: Plural.

Brumbaugh, K. M., & Smit, A. B. (2013). Treating children ages 3–6 who have speech sound disorder: A survey. *Language, Speech, and Hearing Services in Schools, 44*(3), 306–319. https://doi.org/10.1044/0161-1461(2013/12-0029)

Dinnsen, D. A., & Elbert, M. (1984). On the relationship between phonology and learning. In M. Elbert, D. A. Dinnsen, & G. Weismer (Eds.), *Phonological theory and the misarticulating child (ASHA monographs no. 22)* (pp. 59–68). Rockville, MD: American Speech-Language-Hearing Association.

Elbert, M., Dinnsen, D. A., & Powell, T. W. (1984). On the prediction of phonologic generalization learning patterns. *Journal of Speech and Hearing Disorders, 49*(3), 309–317.

Gierut, J. A. (1998). Treatment efficacy: Functional phonological disorders in children. *Journal of Speech, Language, and Hearing Research, 41*(1), S85–S100.

Gierut, J. A. (1999). Syllable onsets: Clusters and adjuncts in acquisition. *Journal of Speech, Language, and Hearing Research, 42*(3), 708–726.

Gierut, J. A. (2001). Complexity in phonological treatment: Clinical factors. *Language, Speech, and Hearing Services in Schools, 32*(4), 229–241. https://doi.org/10.1044/0161-1461(2001/021)

Gierut, J. A. (2007). Phonological complexity and language learnability. *American Journal of Speech-Language Pathology, 16*(1), 6–17. https://doi.org/10.1044/1058-0360(2007/003)

Gierut, J. A., & Champion, A. H. (2001). Syllable onsets II: Three-element clusters in phonological treatment. *Journal of Speech, Language, and Hearing Research, 44*(4), 886–904.

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

Gierut, J. A., & Hulse, L. E. (2010). Evidence-based practice: A matrix for predicting phonological generalization. *Clinical Linguistics & Phonetics, 24*(4–5), 323–334. https://doi.org/10.3109/02699200903532490

Gierut, J. A., & Morrisette, M. L. (2012). Age of word acquisition effects in treatment of children with phonological delays. *Applied Psycholinguistics, 33*(1), 121–144.

Gierut, J. A., Morrisette, M. L., Hughes, M. T., & Rowland, S. (1996). Phonological treatment efficacy and developmental norms. *Language, Speech, and Hearing Services in Schools, 27*(3), 215–230.

Gierut, J. A., & O'Connor, K. M. (2002). Precursors to onset clusters in acquisition. *Journal of Child Language, 29*(3), 495–518.

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

Glaspey, A. M., & Stoel-Gammon, C. (2005). Dynamic assessment in phonological disorders: The scaffolding scale of stimulability. *Topics in Language Disorders, 25*(3), 220–230.

Goldman, R., & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation–Second Edition*. Circles Pines, MN: AGS.

Goldman, R., & Fristoe, M. (2015). *Goldman-Fristoe Test of Articulation–Third Edition*. Circles Pines, MN: AGS.

Kamhi, A. G. (2006). Treatment decisions for children with speech-sound disorders. *Language, Speech, and Hearing Services in Schools, 37*(4), 271–279.

Kirk, C., & Demuth, K. (2005). Asymmetries in the acquisition of word-initial and word-final consonant cluster. *Journal of Child Language, 32,* 709–734.

Levelt, C., Schiller, N., & Levelt, W. (2000). The acquisition of syllable types. *Language Acquisition, 8,* 237–264.

Lleó, C., & Prinz, M. (1996). Consonant clusters in child phonology and the directionality of syllable structure assignment. *Journal of Child Language, 23,* 31–56.

Marshalla, P. (2007). *Fronal lisp, lateral lisp*. Mill Creek, WA: Marshalla Speech and Language.

McLeod, S., van Doorn, J., & Reed, V. A. (2001). Normal acquisition of consonant clusters. *American Journal of Speech-Language Pathology, 10,* 99–110.

McReynolds, L. V., & Jetzke, E. (1986). Articulation generalization of voiced-voiceless sounds in hearing-impaired children. *Journal of Speech and Hearing Disorders, 51,* 348–355.

Miccio, A. W. (2002). Clinical problem solving: Assessment of phonological disorders. *American Journal of Speech-Language Pathology, 11,* 221–225.

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

Piske, T., MacKay, I. R. A., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics, 29,* 191–215.

Powell, T. W. (1991). Planning for phonological generalization: An approach to treatment target selection. *American Journal of Speech-Language Pathology, 1,* 21–27.

Powell, T. W., & Elbert, M. (1984). Generalization following the remediation of early- and later-developing consonant clusters. *Journal of Speech and Hearing Disorders, 49,* 211–218.

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

Powell, T. W., & Miccio, A. W. (1996). Stimulability: A useful clinical tool. *Journal of Communication Disorders, 29*(4), 237–253.

Rvachew, S., & Nowak, M. (2001). The effect of target-selection strategy on phonological learning. *Journal of Speech, Language, and Hearing Research, 44,* 610–623.

Secord, W., Boyce, S. E., Donohue, J. S., Fox, R. A., & Shine, R. E. (2007). *Eliciting sounds: Techniques and strategies for clinicians* (2nd ed.). Clifton Park, NY: Thomson Delmar Learning.

Shriberg, L. D. (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.

Shriberg, L. D., Gruber, F. A., & Kwiatkowski, J. (1994). Developmental phonological disorders. III: Long-term speech-sound normalization. *Journal of Speech and Hearing Research, 37*(5), 1151–1177.

Smit, A. B. (1993a). Phonologic error distributions in the Iowa-Nebraska articulation norms project: Consonant singletons. *Journal of Speech and Hearing Research, 36,* 533–547.

Smit, A. B. (1993b). Phonologic error distributions in the Iowa-Nebraska Articulation Norms Project: Word-initial consonant clusters. *Journal of Speech and Hearing Research, 36,* 931–947.

Smit, A. B., Hand, L., Freilinger, J. J., Bernthal, J. E., & Bird, A. (1990). The Iowa Articulation Norms Project and its Nebraska replication. *Journal of Speech and Hearing Disorders, 55*(4), 779–798.

Templin, M. C. (1957). *Certain language skills in children; their development and interrelationships* (p. 183). Minneapolis, MN: University of Minnesota Press.

Werker, J. F., & Hensch, T. K. (2015). Critical periods in speech perception: New directions. *Annual Review in Psychology, 66,* 173–196. https://doi.org/10.1146/annurev-psych-010814-015104