Children’s acquisition of adult-like speech production has fascinated speech-language pathologists for over a century, and data gained from associated research have informed every aspect of speech-language pathology practice. The acquisition of the consonant cluster has received little attention during this time, even though the consonant cluster is a common feature of speech, its acquisition is one of the most protracted of all aspects of children’s speech development, and the production of consonant clusters is one of the most common difficulties for children with speech impairment. This paper reviews the literature from the past 70 years to describe children’s normal acquisition of consonant clusters. Articulatory, phonological, linguistic, and acoustic approaches to the development of consonant clusters are reviewed. Data from English are supplemented with examples from other languages. Consideration of the information on consonant cluster development revealed 10 aspects of normal development that can be used in speech-language pathologists’ assessment and analysis of children’s speech.

Key Words: normal development, consonant clusters, phonology, children’s speech

Consonant clusters are a feature of many of the world’s languages. In a study of 104 world languages, based on the work of Greenberg (1978), Locke (1983) calculated that 39% had word-initial clusters only, 13% had final clusters only, and the remaining 48% had clusters in both word-initial and word-final position. In English, one third of monosyllables begin with a consonant cluster, and consonant clusters predominate in word-final position (Locke, 1983). This predominance in word-final position is attributable to the addition of the phonemes /s, z, t, d/ to indicate grammatical morphemes. When such morphophonemic clusters are excluded, only 18% of English monosyllables end in consonant clusters. Some languages (such as Italian) have more consonant clusters than English does, and other languages (such as Cantonese, Catalan, Portuguese, and Turkish) have fewer (Greenberg, 1978; Swan & Smith, 1987). For example, Cantonese has only two consonant clusters, namely word-initial /kw/ and /kʰw/ (So & Dodd, 1995).

Children learning to produce consonant clusters in any language have a challenging task, and those learning English have a uniquely complex situation. The large variety of clusters permissible in English, both at the beginning and at the end of syllables, makes even monosyllables extraordinarily complex (in words such as *strength*). A further complicating factor is that morphological endings create even more complex phoneme sequences (e.g., *sixths*). Consequently, the acquisition of clusters is one of the longest-lasting aspects of speech acquisition in normally developing children. Children as young as 2 years of age produce some consonant clusters correctly (Preisser, Hodson, & Paden, 1988; Stoel-Gammon, 1987; Watson & Scukanec, 1997a). Yet some 8- to 9-year-olds are still mastering consonant clusters (Smit, Hand, Freilinger, Bernthal, & Bird, 1990; Templin, 1957). Because of the protracted acquisition of consonant clusters, gradual developmental gains can be identified and described in greater detail than in the acquisition of singleton consonants, which is essentially completed much earlier (Smit et al., 1990).

This article provides a tutorial about children’s normal acquisition of consonant clusters. Knowing about consonant clusters and their development is important for speech-language pathology practice because it can help to determine whether children’s speech development is progressing normally and can assist in selecting targets for intervention. The majority of pediatric clients with speech impairment are reported to have difficulty producing adult-like consonant clusters (Andrews & Fey, 1986; Baker, 2000; Beers, 1993; Chin & Dinsen, 1992; Crary, 1983; Dodd & Iacano, 1989; Elbert, Dinsen, & Powell, 1984; Garn-Nunn, 1986; Grunwell, 1987; Hodson, 1982; Hodson & Paden, 1981; J. C. L. Ingram, 1989; Khan, 1982; Leahy
Consonant Cluster Production
From Birth to Mastery

Children’s first utterances and first words are assumed not to contain consonant clusters. However, no research specifically confirms this assumption. Consonant clusters are not mentioned in the literature describing the gurgling, cooing, or babbling of children under 1 year of age (e.g., Mitchell & Kent, 1990; Robb & Bleile, 1994; Stark, Bernstein, & Demorest, 1993; Stoel-Gammon & Cooper, 1984), nor are consonant clusters mentioned in the literature that describes children up to the age when they produce their first 50 words (e.g., Donahue, 1986; Ferguson & Farwell, 1975; Leonard, Newhoff, & Mesalam, 1980; Robb & Bleile, 1994; Stark et al., 1993; Stoel-Gammon, 1985; Stoel-Gammon & Cooper, 1984). For example, Robb and Bleile (1994) conducted a comprehensive longitudinal study of the acquisition of consonant inventories of seven children between the ages of 8 and 25 months, but did not mention consonant clusters.

The ability to produce consonant clusters is reported to emerge when children are around 2 years of age (e.g., French, 1989; Lleo & Prinz, 1996) during the phase that Ingram (1991) refers to as the “word spurt.” French’s (1989) son (who was described as a late talker) produced his first word containing a consonant cluster at age 1;10 (years, months), when he said [bv] for /br/. At age 1;11, he produced [nj], and by age 2;2 he had made five other attempts at target clusters, none of which was correct. Lleo and Prinz (1996) reported that children’s correct production of consonant clusters in German and Spanish began at age 1;10 for word-initial clusters and as early as age 1;5 for medial clusters.

Several reasons have been proposed for the emergence of children’s ability to produce consonant clusters in the second year of life. Ingram (1991) suggests that the “word spurt” may be linked to a significant development in children’s phonological analysis of the receptive vocabulary in terms of phonotactics. Consonant clusters represent an important departure in phonotactics from the earlier word shapes of CV, VC, or CVC. The ability to produce consonant clusters may also be related to maturation of the children’s motor speech mechanism and continued anatomical development of the oromusculature.

The following sections review what is known about children’s early attempts at consonant cluster production and describe those attempts at consonant clusters that do not match adult forms. An overview of information from studies that explain children’s mastery of various consonant clusters is also presented.

Earliest Attempts at the Production of Consonant Clusters

From 2 years of age, children begin to produce two and sometimes three consonants together within a word. However, these early attempts at the production of consonant clusters usually result in the production of forms that are inconsistent with the ambient language. Reports of children’s earliest attempts at consonant cluster production will be examined initially by considering the children’s productions independently from the adult target, using their inventory of consonant clusters, word shapes, and the number of different consonant clusters produced. Next, the productions will be reviewed by comparing them with the adult targets, looking at the nature of the differences between the children’s non-adult forms and the adult target with respect to mismatches, phonological processes, description of the deleted member, acoustic analyses, and homonymy. Finally, mastery of consonant clusters will be discussed.

Inventory of Consonant Clusters. The independent inventories of consonant clusters of 2- to 3-year-old children are frequently reported to contain word-initial consonant clusters that are not permitted in the ambient language. For example, Dyson (1988) studied the phonetic inventories of ten 2- and 3-year-old children and found that the only word-initial cluster used by over half of the subjects was [fw]; consonant clusters used less frequently included [bw, kw, tr, sp, st, sn, sl]. Watson and Scukanec (1997b) reported that the type of cluster produced changed over time, from the labial clusters [pw] and [bw] that were not present in the ambient language, to the permissible clusters [st, sp], and [pl] by age 3;0.
The word-final independent inventories of consonant clusters of 2- to 3-year-olds often contain more consonant clusters than the word-initial inventories, and those clusters are generally permissible in English. The impact of morphological development (e.g., the emergence of the plural and past tense morphemes) may result in an increase in the number of occurrences of word-final clusters. It may also provide a rationale for instability in the accuracy of productions of word-final clusters as children grapple with competing morphological and phonological demands (Crystal, 1987).

McLeod, van Doorn, and Reed (in press-a) stated that the most common word-final consonant clusters produced by the 2-year-old subjects in their study contained nasals and are the most common word-final clusters produced by the children in their data set. Other researchers who have included morphophonemic clusters in the repertoire, such as Watson and Scukanec (1997b), reported production of [-nd], [-nt], [-nk]); however, morphophonemic clusters were excluded from their data set. According to Dobrich and Scarborough (1992), between the ages of 2;0 and 5;0 children attempt target words containing word-initial consonant clusters to a degree similar to adults. For their subjects, word-final consonant clusters were produced to a lesser degree at age 2;0 but corresponded to adult production by age 3;0. These authors suggested that word-final consonant clusters were limited at 2 years of age because morphophonemic clusters created by the addition of plural and possessive morphemes may not have been established.

Non-Adult Productions of Consonant Clusters

Although children as young as 2 years of age can produce consonant clusters, typically their attempts result in non-adult productions—either a reduction in the number of elements of the consonant cluster, production of differing phonemes with the retention of the correct syllable shape, or changes in both the syllable shape and constituent phones. The importance of the study of non-adult productions was highlighted by Prather, Hedrick, and Kern (1975) in the conclusion of their paper on the acquisition of phonemes:

It seems probable that error patterns in normally developing children represent predictable changes… Perhaps an important prognostic indicator of articulation deviancy is evident in the error patterns even at the age of two or three years. (p. 190)

A number of approaches have been applied to the consideration of non-adult productions of consonant clusters. These include: mismatches, phonological processes, describing the deleted member, acoustic analysis, and homonymy. Each of these will be considered in turn.

Mismatches. Smit (1993) tabulated by age range and frequency the errors made on word-initial consonant clusters by subjects in the Smit et al. (1990) study. Smit’s (1993) results were comprehensive, providing the percentage of specific errors for each age and each word-initial consonant cluster. For example, she reported that 2- to 3-year-olds produced /bl/ in block correctly 30% of the time, as [b] 15–50% of the time, and as [l] less than 3% of the time. Smit (1993) provided a summary of the typical mismatches. She found that typical errors were (a) reduction to the obstruent in obstruent + approximant clusters and (b) reduction to the second element in /l/ clusters. When clusters were preserved but with one element in error, the error was typically the same as for the singleton consonant.

Phonological Processes. One of the more popular ways to describe the development of consonant clusters has been to define which phonological processes are operating. The theory of Natural Phonology (Stampe, 1969) formed the basis of the phonological process approach and transformed the way that children’s speech sound errors were viewed in the 1980s (e.g., Grunwell, 1985a, 1985b; Hodson, 1986; Ingram, 1981; Shriberg & Kwiatkowski, 1980). A number of phonological processes can be applied to consonant cluster production, the most common being cluster reduction. Others include cluster simplification, epenthesis, coalescence, and metathesis.
Cluster reduction is defined as “the deletion of one or more consonants from a target cluster so that only a single consonant occurs at syllable margins” (Grunwell, 1987, p. 217); in the case of clusters containing three elements, either one or two consonants may occur at syllable margins. Cluster reduction has been described as “the most common and longest lasting stage” in the development of cluster production (Shirberg & Kwiatkowski, 1980, p. 138) and has frequently been described in the speech of normally developing children (e.g., Dodd, 1995; Dyson & Paden, 1983; Grunwell, 1981; Haelsg & Madison, 1986; Hodson, 1982; D. Ingram, 1989; Khan, 1982; Locke, 1983; McCormack & Knighton, 1996; McLeod, 1999; Olmsted, 1971; Preisser et al., 1988; Roberts, Burchinal, & Footo, 1990; Shirberg & Kwiatkowski, 1980; Watson & Scukanec, 1997a). In fact, cluster reduction occurs more frequently than most if not all other phonological processes (whether for singleton or cluster targets), especially in younger subjects (Dyson & Paden, 1983; Preisser et al., 1988; Roberts et al., 1990; Watson & Scukanec, 1997a). As children become older, the occurrence of cluster reduction diminishes (Roberts et al., 1990). An example of the dramatic decrease in the occurrence of cluster reduction is presented by Preisser et al. (1988), who reported that the mean occurrence of cluster reduction fell from 93% for children aged 1:6 to 1:9 to 51% for children aged 2:2 to 2:5. Table 1 provides the reported percent occurrence of cluster reduction for normally developing English-speaking children of various ages.

The presence of cluster reduction has not only been documented for English-speaking children. Cluster reduction also occurs commonly across languages other than English that contain consonant clusters (Dutch—Beers, 1993; Danish—Bloch, 1996; Italian—Bortolini & Leonard, 1991; Telugu—Chervela, 1981; German—Fox & Dodd, 1999; German and Spanish—Lleo & Prinz, 1996; Cantonese—So & Dodd, 1995; Portuguese—Yavas & Lamprecht, 1988; Turkish—Kopkalli-Yavuz & Topbas, 1998). For example, Bortolini and Leonard (1991) studied 2-year-old Italian children. They found that all of nine normally developing children demonstrated liquid cluster reduction, five demonstrated sibilant cluster reduction, and two demonstrated nasal cluster reduction. Yavas and Lamprecht (1988) found that in four Portuguese-speaking children from Brazil, aged 7 to 9 years, “subjects confirm" the widespread occurrence of cluster reduction (the only absolutely exceptionless process in our data)” (p. 339).

Cluster simplification occurs when two elements of the cluster are produced, but one or both of the elements are produced in a non-adult manner. The phonological processes that effect systemic simplifications of singleton phonemes (such as gliding, stopping, and fronting) are often described as cluster simplification in the context of a consonant cluster. The most commonly reported instance of cluster simplification results in the gliding of approximants (i.e., /w, r, l, j/) (e.g., Long, Fey, & Channell, 1998). Usually, the liquids /l, r/ are realized as [w, j] (Grunwell, 1987). For example, gliding occurs when the word green /gri:n/ is produced as [gwin]. Cluster simplification has been reported to occur in sequence with cluster reduction (Smit, 1993; Watson & Scukanec, 1997a). For example, Watson and Scukanec (1997a) reported that cluster simplification showed a pattern of increase followed by decline in their 12 subjects, aged 2:0 to 3:0 (see Table 2). The increase in the occurrence of cluster simplification seems to coincide with a decrease in the occurrence of cluster reduction. The interrelationship between the acquisition of the correct number of elements in the cluster and the refinement of all the cluster elements is important. For example, in the Smit (1993) data, a reason for the length of time to produce three-element consonant clusters may be the lag in modifying substitutions for /ls/ and /rs/, well after the appropriate number of elements of the cluster has been achieved.

Epenthesis is the insertion of a vowel (frequently schwa) between the consonants within the cluster, and as such effects a change in the syllable shape (Shirberg & Kwiatkowski, 1980). For example, plate /plet/ becomes [plet]. Epenthesis has been reported in the speech of 2- to 3-year-olds (Bortolini & Leonard, 1991; Dyson & Paden, 1983; Higgs, 1968; McLeod, van Doorn, & Reed, 2001).

### TABLE 1. Percent occurrence of cluster reduction in English-speaking children.

<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>1:6</th>
<th>2:0</th>
<th>2:3</th>
<th>2:6</th>
<th>2:9</th>
<th>3:0</th>
<th>3:6</th>
<th>4:0</th>
<th>4:6</th>
<th>5:0</th>
<th>6:0</th>
<th>7:0</th>
<th>8:0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haelsg &amp; Madison (1986)</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>30</td>
<td>18</td>
<td>10</td>
<td>15</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McCormack &amp; Knighton (1996)</td>
<td>22</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McCormack &amp; Knighton (1996)</td>
<td>28</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preisser et al. (1988)</td>
<td>60</td>
<td></td>
<td>93</td>
<td>76</td>
<td>51</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roberts et al. (1990)</td>
<td>145</td>
<td></td>
<td>68</td>
<td></td>
<td></td>
<td>42</td>
<td>25</td>
<td>15</td>
<td></td>
<td>10</td>
<td>7</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Watson &amp; Scukanec (1997a)</td>
<td>12</td>
<td></td>
<td></td>
<td>46</td>
<td>48</td>
<td>34</td>
<td>25</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** Blank spaces indicate that children of that age group were not studied. F = female, M = male.

*a*Preisser et al. (1988) used age ranges 1:6–1:9, 1:10–2:1, and 2:2–2:5.

Epenthesis is the insertion of a vowel (frequently schwa) between the consonants within the cluster, and as such effects a change in the syllable shape (Shirberg & Kwiatkowski, 1980). For example, plate /plet/ becomes [plet]. Epenthesis has been reported in the speech of 2- to 3-year-olds (Bortolini & Leonard, 1991; Dyson & Paden, 1983; Higgs, 1968; McLeod, van Doorn, & Reed, 2001).
in press) as well as in older children (Ingram, Pittam, & Newman, 1985; Olmsted, 1971). For example, Ingram et al. (1985) reported that epenthesis (which they called “breaking”) occurred frequently in children in Grade 2 and to a lesser extent in children in Grades 3 and 4. Epenthesis also may occur when foreign words are borrowed that contain consonant clusters not permissible in the native language. For example, the Turkish version of the French word for beach (plage) is pronounced [pi laz].

Coalescence occurs when the reduced cluster contains a new consonant composed of features from the original consonants. Thus, swim /swim/ becomes [frm] because the [+ fricative] feature of /s/ co-occurs with the [+ labial] feature of /wl/, resulting in a labial fricative, [f]. Coalescence has been reported to occur in the speech of 2- to 3-year-olds by Dyson and Paden (1983).

Metathesis is the reversal of adjacent segments or migration of an element within the word (Shriberg & Kwiatkowski, 1980). For example, ask is produced as [aks] instead of /ask/. Metathesis has occurred in the speech of young children (Bortolini & Leonard, 1991; Edwards & Shriberg, 1983; McLeod et al., in press-b; Stockman & Stephenson, 1981). However, as Olmsted (1971, p. 247) states, “the incidence (in clusters) of metathesis is negligible.”

Describing the Deleted Member. Many researchers who have described normal consonant cluster development have presented realization rules about the element of the consonant cluster that will be deleted if cluster reduction occurs (e.g., Dodd, 1995; Gierut, 1999; Greenlee, 1973; Grunwell, 1987; Higgs, 1968; Preisser et al., 1988; Smit, 1993). A number of the original theoretical discussions around the deletion of an element of a consonant cluster appealed to markedness theory (Toombs, Singh, & Hayden, 1981), and distinctive features were used to describe the aspects that were marked. Locke (1983) conducted a crosslinguistic comparison of occurrences of the deleted members of consonant clusters. His comparison of various studies led him to conclude that there were few exceptions to the following summary:

If there is a glide or a liquid present, it typically will be the second member, and children will omit. In most other cases, the first member will be a stop or a fricative, and children will omit the stop or fricative. If both members are stops, fricatives, or nasals, the first stop, fricative, or nasal will be omitted. (Locke, 1983, p. 71)

Recently, research has indicated that the sonority sequencing principle may provide a theoretical basis for the description of the deleted member of the consonant cluster (Chin, 1996; Gierut, 1999; Ohala, 1999). The sonority sequencing principle is a “presumed universal which governs the permissible sequence of consonants within syllables” (Gierut, 1999, p. 708). Phonemes with low sonority values are found at syllable margins, and phonemes with higher sonority values are located towards the center of the syllable (Clements, 1990) (see Wyllie-Smith & McLeod, 2001). For example, Ohala (1999) found that normally developing children reduce word-initial consonant clusters in a manner that produces a maximal rise in sonority. This pattern of reduction uses the sonority hierarchy, which “predicts differential production of the same consonant depending on the type of cluster within which the consonant is contained” (Ohala, 1999, p. 402).

Acoustic Analysis of Consonant Cluster Development. Acoustic analyses, particularly using the spectrograph, have been conducted to describe subtle and sometimes auditorily undetectable features of speech production. The literature contains a few reports of acoustic characteristics of reduced clusters where children have substituted stops for /s/ + stop word-initial clusters (normal speech development—Bond & Wilson, 1980; Catts & Kamhi, 1984; McLeod, van Doorn, & Reed, 1996; impaired speech development—Scobbie, 1995; Smit & Bernthal, 1983; Tyler, 1995). For example, Catts and Kamhi (1984) performed a longitudinal study with measurements for six children, initially aged 1;9 to 2;10 years, over a period of 5 to 17 months, until correct cluster production was achieved. The researchers found that the children consistently produced short-lag stops as substitutes for clusters. In contrast, in a cross-sectional study of a group of five children aged 1;10 to 3;0 years, Bond and Wilson (1980) found that two children used long-lag voice onset time (VOT) for the substituted stop, one child used short-lag VOT, and the other two used both. Neither of these studies made a direct comparison of VOT for the singleton and substituted cluster contexts to establish whether the VOT for long-lag stops used as a cluster substitute was shorter (but still in the long-lag range) than that of stops in the singleton context. McLeod et al. (1996) studied sixteen 2-year-olds and found that for word-initial /s/ + stop clusters, that had been reduced to a stop, the VOT for the stop in the cluster target word was significantly less than that for the singleton target word.

Temporal data for the realization of consonant clusters in children’s speech generally reflect the same trends as adult data (see Weismer, 1984, for a review): shorter VOT for stops in /s/ + stop context (Scobbie, 1995), shorter duration for /s/ as the first element in two-element clusters (Gilbert & Purves, 1977; Hawkins, 1979; Weismer, 1984; Weismer & Elbert, 1982), and longer VOT for voiceless stops in stop + /l, r/ clusters (Menyuk & Klatt, 1975). Whenever children’s speech is compared with adult speech, there are universal findings of longer segment durations with increased variability (Kent, 1976; Smith, 1978; Weismer, 1984). There has been some debate about the nature of the variability that is characteristic of children’s speech: whether it is a mathematical consequence of slower speech with longer segment durations or whether it is an independent consequence of a less mature speech motor system (Smith, 1994). On balance, there is evidence to suggest that the degree of variability cannot be entirely accounted for by the longer segment durations found in children’s speech, but is a direct consequence of a more variable speech motor production system (Kent, 1992; Smith, 1994).

Homonymy. A final analysis of the non-adult realizations of consonant clusters is to consider the production of homonyms that result in a breakdown of communicative
clarity. A homonym occurs when phonological contrasts are neutralized and the resulting production of a particular word is not auditibly different from another word. For example, if a child attempts to produce *snail* and instead produces *[nail]*, this word is a homonym with the word *nail*. Homonyms are common in normal language acquisition (Ingram, 1975, 1985; Lleo, 1990; Locke, 1979; Priestly, 1980; Stoel-Gammon & Cooper, 1984; see Vihman, 1981, for a review). Leinonen-Davies (1988) hypothesized that homonymy would occur in normally developing children’s speech as a result of cluster reduction. McLeod, van Doorn, and Reed (1998) described the nature and occurrence of homonymy in normally developing 2- to 3-year-old children’s productions of consonant clusters. Homonyms occurred in approximately one-fifth of the task items. There was a reduction in the occurrence of homonymy as the subjects neared their third birthdays. The 2-year-old subjects displayed more instances of homonymy as a result of cluster reduction (e.g., *[ki]* for *ski* and *key*), whereas the 3-year-old subjects tended to have more instances of homonymy as a result of cluster creation (e.g., *[snail]* for *snail* and *sail*).

**Mastery of Consonant Clusters**

The mastery of consonant clusters is a protracted process. As mentioned earlier, children as young as 2 years of age have been reported to produce adult-like consonant clusters. Shriberg and Kwiatkowski (1980) reported that 4-year-olds correctly produced 90% of consonant clusters in spontaneous speech. In contrast, Smit et al. (1990) reported that very few clusters had been mastered by 4 years of age, with the majority being mastered at age 6 or 7 years, and the last clusters not mastered until age 8 to 9 years (see Table 3).

Two major studies have been undertaken that present age of acquisition findings for consonant clusters. Templin (1957) studied word-initial and word-final clusters, and Smit et al. (1990) studied word-initial consonant clusters. Table 3 presents a comparison of age-of-acquisition findings.

According to Smit et al. (1990), 75% of normally developing children produce a consonant cluster consisting of a stop + /w/ (e.g., *queen*) by age 3,6; clusters containing /l/ (e.g., *play*) excluding /sl/ by age 4;6 to 5;6; clusters containing /l/ (e.g., *train*) excluding /θl/ by age 6;0; clusters containing /s/ (e.g., *sweep, stop*) and /θl/ (through) by age 7;0. Despite differences in methodology and an intervening period of 30 years, the studies of Templin (1957) and Smit et al. (1990) show many similarities in their overall findings but differences in specific findings. For example, each found that the earliest word-initial consonant clusters to be produced correctly were /tw/ and /kw/. Similar consensus was reached that among the most difficult consonant clusters were /skl/ and /spr/; however, the age of acquisition for Templin’s (1957) subjects was one year younger than for those of Smit et al. (1990).

Several smaller studies have been conducted that also have considered age of acquisition of specific consonant clusters. Subjects generally achieved accuracy later than did the subjects described by Templin (1957). Arlt and Goodban (1976) assessed six consonant clusters, but did not report age of acquisition findings within the text of their article. However, they did report in a table caption that /gr/ and /br/ were produced 6 months later by their subjects than by Templin’s (1957) subjects. Higgs (1968) studied the age of acquisition for the consonant clusters /spl, /st/, and /skl/ (see Table 3). She reported a steady increase in the percentage of consonant clusters correct from ages 2;6 to 5;0. At age 2;6, the percentages of consonant clusters correct for /spl, /st/, and /skl/ were 38, 38, and 37, respectively; and at age 5;0, they were 84, 84, and 84. These results from Higgs (1968) were consistently lower than those obtained by Templin (1957), indicating that Templin’s subjects developed the clusters earlier than did Higgs’ subjects. Higgs’ results also differ from those of Paynter and Petty (1974), who found that 57% of 2.5-year-old girls produced /st/ correctly, but fewer than 50% of boys could produce the cluster.

Some consonant clusters are easier to master than others. Children typically master consonant clusters that consist of stop + liquid elements (e.g., /pl/) before they master fricative + liquid clusters (e.g., /sl/). Ingram (1976; Smit et al., 1990; Smith, 1973; Templin, 1957; Wellman et al., 1931). For example, Powell and Elbert (1984) noted that 75% of Templin’s (1957) 4-year-old subjects were able to produce all stop + liquid clusters (except /gr/), but they could not produce any fricative + liquid clusters. This finding was supported by Powell (1993), who studied 4- and 5-year-old children to determine the factors that accounted for variance in the production of consonant clusters. Eleven factors were identified. The most prominent was the presence of a liquid segment in a cluster, which accounted for 42.8% of the variance. Other factors included Factor II—word-final with alveolar fricative, III—word-initial /s/ clusters, IV—word-initial /f/ clusters, V—three-element clusters, and VI—word-final /l/ clusters. Powell (1993) also reported that the position in which the cluster occurred (i.e., word-initial versus word-final) was

---

**Table 3. Age of acquisition for word-initial consonant clusters (75% criterion).**

<table>
<thead>
<tr>
<th>Clusters</th>
<th>Smit et al. (1990)</th>
<th>Templin (1957)</th>
<th>Higgs (1968)</th>
</tr>
</thead>
<tbody>
<tr>
<td>/tw, kw/</td>
<td>3;6</td>
<td>3;6</td>
<td>4;0</td>
</tr>
<tr>
<td>/sp, st, sk/</td>
<td>4;6</td>
<td>5;0–6;0;*</td>
<td>4;0–4;6</td>
</tr>
<tr>
<td>/sm, sn/</td>
<td>5;6</td>
<td>5;0–7;0</td>
<td>4;0</td>
</tr>
<tr>
<td>/sw/</td>
<td>4;6*</td>
<td>6;0</td>
<td>7;0</td>
</tr>
<tr>
<td>/sl/</td>
<td>6;0</td>
<td>7;0</td>
<td>7;0</td>
</tr>
<tr>
<td>/pl, bl, kl, gl, fl/</td>
<td>4;0–4;6</td>
<td>4;0–5;6</td>
<td>4;0–5;0</td>
</tr>
<tr>
<td>/pr, br, tr, dr, kr, gr, fr/</td>
<td>4;6–6;0</td>
<td>5;0–6;0</td>
<td>4;0–4;6</td>
</tr>
<tr>
<td>/θl/</td>
<td>7;0</td>
<td>7;0</td>
<td>7;0</td>
</tr>
<tr>
<td>/skw/</td>
<td>4;6*</td>
<td>7;0</td>
<td>6;0</td>
</tr>
<tr>
<td>/spl/</td>
<td>6;0</td>
<td>7;0</td>
<td>7;0</td>
</tr>
<tr>
<td>/spr, str, skt/</td>
<td>8;0</td>
<td>8;0</td>
<td>5;0–7;0</td>
</tr>
</tbody>
</table>

* A reversal occurs in older age groups.
not a factor in the difficulty of the cluster for the 4- to 5-year-olds, but that it may be for younger children. Three-element clusters were more difficult to produce than two-element clusters.

There are inconclusive results regarding whether correct production of consonant clusters occurs at a later stage than correct production of the singleton components of clusters. Templin’s (1957) ages of acquisition for clusters were generally earlier than her ages for the singleton elements that made up those clusters. In contrast, the findings of Smit et al. (1990) showed that later-developing singleton components may be learned later than or simultaneously with the consonant clusters of which they are a part. For example, girls achieved a 75% level of acquisition for /l/ by age 4;6 and for /ll/ clusters by age 4;0 to 6;0. In fact, three clusters, /pl, bl, kl/, achieved a 75% level of acquisition 6 months earlier than the singleton /l/ phoneme.

Developmental Progression of Consonantal Cluster Acquisition

One of the problems of traditional assessment and analysis is that speech development is regarded as the achievement of correct pronunciation (including description of non-adult productions), and there is no description of the gradual processes of phonological learning and articulatory mastery and maturation. The approaches mentioned thus far for describing the acquisition of phonology conceal the progressive stages through which a child may pass. As Elbert (1984) suggested, developmental norms have provided an outline of the acquisition process without showing the details. A further approach to understanding the normal acquisition is to consider the developmental sequence undertaken by children to achieve mastery of consonant clusters (e.g., Hutcheson, 1968; Sander, 1972). Hall, Adams, Hesketh, and Nightingale (1998) called for a measure of “degrees of change” in a child’s performance in order to demonstrate progressive approximations to the adult target. There have been a number of approaches for describing the developmental progression in the acquisition of consonant clusters. The two main approaches are graphical representation of development and developmental stages.

Graphical Representation of Development

A developmental approach to children’s speech acquisition was presented pictorially by Sander (1972). Instead of stating actual ages at which children acquire a phoneme, Sander suggested that children’s data be presented graphically to indicate the gradual acquisition over time. Also adopting a graphing approach, Grunwell (1981) presented the age of suppression of phonological processes, including cluster reduction, in a graph that indicated that suppression of the processes was gradual. Smit et al. (1990) presented their age of acquisition data as a line graph, indicating the percentage of subjects who could produce each phoneme at each age. McLeod et al. (in press-a) graphically presented longitudinal data from 16 normally developing 2-year-old children to demonstrate the gradual acquisition of consonant clusters, as well as the interaction between the use of cluster reduction and cluster simplification.

Once data are considered graphically as a sequence of development rather than as an age at which consonant clusters are mastered, it is apparent that the path to consonant cluster acquisition is not one of steady progression. Many researchers have documented frequent reversals in age of acquisition data (Dyson & Paden, 1983; Smit et al., 1990; Templin, 1957). For example, Smit (personal correspondence, July 1994) reported that when their male subjects produced /s/ + nasal clusters, they had 73% acceptable productions at age 3;6, 59% at ages 4;0 and 4;6, 72% at age 5;0, another reversal to 68% at age 5;6, 69% at age 6;0, and a steady increase to age 9;0. Reversals and revisions have been reported throughout the speech and language acquisition literature. Dyson and Paden (1983, p. 16) said that “steady ‘improvement’ toward the adult model was an exception.” Reversals have been reported for general phonological acquisition (Donahue, 1986; Dyson & Paden, 1983) as well as specific aspects, including the acquisition of singleton /s/ targets (Poole, 1934; Prather et al., 1975; Templin, 1957). For example, Donahue (1986, p. 212) noted that between ages 1;3 and 1;6 her son demonstrated a “U shaped developmental curve” in the use of two-word utterances as a result of his reliance on a consonant harmony rule.

Developmental Stages

Despite the finding that children’s sequence of development is typically marked by reversals and revisions, a number of researchers have hypothesized that there are common stages in the development of consonant clusters. For example, 30 years ago Higgs stated, “the phonetic acquisition of these initial clusters and perhaps of speech generally is best viewed as a step-by-step development” (Higgs, 1968, p. 138). One person has had a significant impact on the description of these stages of development of consonant clusters. Greenlee (1974) proposed a three-stage route to the development of stop + liquid consonant clusters: (1) liquid deletion, (2) substitutions for the adult liquid, and (3) correct production. She stated that the stages were confirmed by the data from the six languages she studied, but that not all words gave evidence for each stage. She also described several other “subprocesses,” for example, “deletion of the entire sequence” and “velar-dental interchange (consonant harmony).” Greenlee’s stages were summarized and expanded by Elbert and McReynolds (1979) to include all two-element clusters:

1. Both segments are omitted (e.g., blue produced as [u])
2. One segment of the cluster is used while the other is omitted (e.g., blue produced as [bu])
3. Both segments are marked in some way (e.g., blue produced as [bwu])
4. Both segments are used appropriately (e.g., blue produced as [bul])

Elbert and McReynolds’ version of Greenlee’s stages has been found to be a robust description of children’s...
development of two-element consonant clusters by a number of other researchers (Chin & Dinnisen, 1992; Elbert & McReynolds, 1979; McLeod, van Doorn, & Reed, 1997, in press-b; Smit, 1993). Despite the robustness of these descriptions, there is some evidence that not all children pass through each of these stages for the acquisition of each consonant cluster. Consequently, Dyson and Paden (1983) suggested three “ordered sets” that children may adopt for the acquisition of consonant clusters:

1. Cluster reduction → correct
2. Cluster reduction → one segment distorted or substituted → correct
3. Deletion of final cluster → cluster reduction → correct final cluster

Researchers of languages other than English have described a similar developmental route in the acquisition of consonant clusters. For example, Chervela (1981) studied four children aged 1;6 to 3;0 who spoke Telugu, a language of India. Chervela suggested the presence of the following hierarchy among the processes during the children’s acquisition of consonant clusters: total deletion → reduction + substitution → substitution + assimilation → substitution of one or both consonants → adult cluster. Chervela added that total deletion occurred only in the earliest stage of development, reduction + substitution occurred more for word-initial clusters, and substitution + assimilation occurred more for medial clusters.

In a different approach to the consideration of developmental stages, Lleo and Prinz (1996) studied the acquisition of consonant clusters in German- and Spanish-speaking children. They found marked differences between the German- and Spanish-speaking children’s acquisition strategies, but described a general developmental sequence of CV → CVC → CVCC → CCVCC. This developmental view coincided with those of Chervela (1981) and Greenlee (1974) as it described the increase in complexity of the syllable shape; however, Lleo and Prinz added a second dimension by suggesting that consonant clusters occurred in word-final position before the word-initial position.

Taking a broad perspective, children seem to move through a similar progression in learning consonant clusters (Elbert & Gierut, 1986): from a one-element realization to a two-element realization and finally to a correct realization. Current knowledge, however, is insufficient to describe the significance or occurrence of total deletion of all elements in the acquisition of consonant clusters.

**Individual Variability**

Although overall there appears to be a general developmental route in the acquisition of consonant clusters, many researchers have identified wide individual variation (Dyson & Paden, 1983; Grunwell, 1981; McLeod, 1999; Vihman & Greenlee, 1987; Watson & Scukanec, 1997b). Phonological development is gradual, and variability occurs as gradual change toward mastery. For example, Watson and Scukanec (1997b) studied 12 children every 3 months between the ages of 2;0 and 3;0 years. They noted “great variation in the consonant clusters produced by the subjects” (p. 7). Similarly, Dyson and Paden (1983, p. 16) described the period of 2 to 3 years of age to be “one of extreme variability with subjects ‘trying out’ a variety of strategies to approximate the adult model.” Unfortunately, examination of specific consonant clusters in individual inventories is rarely possible, because data are frequently grouped (e.g., Dyson & Paden, 1983). Two exceptions are Smith (1973), who presented an extensive appendix of the productions of words spoken by his son, and McLeod et al. (in press-a), who presented individualized longitudinal data for the production of consonant clusters by 16 normally developing 2-year-olds. There is need for further examination of variability in children’s speech. Haynes (1998, p. 387) stated, “The research into individual differences is only in its infancy. There are not many reports of such differences in the literature.”

**Characteristic Features of Normal Consonant Cluster Acquisition**

The acquisition of consonant clusters is a protracted process that begins with attempts to produce two adjacent phonemes within a syllable. Even at 2 years of age, some form of consonant cluster can be produced. However, few early attempts result in adult-like productions of consonant clusters. More typically, non-adult forms are produced. These attempts frequently result in the reduction of the number of elements of the consonant cluster and the production of a homonym. There have been many studies using linguistic, phonological, and acoustic analyses that describe features of the reduced elements. A number of other realizations (apart from the reduction of elements) is also apparent, affecting both the syllable shape and the constituent phones. Despite a developmental route marked by reversals and revisions and extreme individual variability, eventually mastery of each consonant cluster occurs.

A comprehensive understanding of consonant cluster development can provide a snapshot of children’s speech systems due to the interrelationship between segmental and syllabic structures. However, this opportunity for gaining an overview of children’s speech systems is often ignored. Typically, during assessment and analysis of consonant cluster production, speech-language pathologists consider the age of acquisition norms for consonant clusters and describe the percentage of occurrence of cluster reduction. Inconsistencies in age of acquisition norms, the advances in analysis techniques, and evidence of heterogeneity within the population for the development of consonant clusters provide impetus for moving beyond a restricted view of consonant cluster development to a broader understanding of interwoven intricacies of the acquisition of consonant clusters. This broadening of focus supports the current trend to acknowledge the heterogeneity among individuals and the need for a variety of theoretical approaches (Powell, Elbert, Miccio, Strike-Roussos, & Brasseur, 1998).

The authors have generated a list of general trends found in 70 years of literature on the normal development of consonant clusters. This list can assist in the assessment,
analysis, and selection of appropriate intervention targets for children with phonological impairment. Observations of the following phenomena may indicate normal development of consonant clusters in children:

1. Two-year-old children can produce consonant clusters, but these clusters may not be of the same form as the ambient language.

2. Word-final consonant clusters generally appear in inventories earlier than word-initial clusters. Children’s production of word-final consonant clusters is increased by the emergence of grammatical morphemes (e.g., plurals and past tense) and consequently the creation of morphophonological consonant clusters (e.g., [-ts] as in cats).

3. Two-element consonant clusters are generally produced and mastered earlier than three-element clusters. There is inconclusive evidence regarding whether phonemes are mastered in singleton contexts before they can be accurately produced in clustered contexts.

4. Consonant clusters containing stops (e.g., /pl/, /kw/) are acquired generally before consonant clusters containing fricatives (e.g., /st/, /dr/).

5. Young children typically delete one element of a consonant cluster (cluster reduction), and this deletion may be explained by principles of markedness and sonority.

6. Homonymy occurs in young children’s attempts to produce consonant clusters. Homonymy frequently occurs as a result of cluster reduction; however, homonyms can also occur as a result of cluster creation.

7. There are a number of other non-adult realizations of consonant clusters; the most common is cluster simplification, with others including epenthesis and coalescence. Metathesis is rare.

8. The acquisition of consonant clusters is gradual, and there is a typical developmental sequence. It is not an all-or-nothing process. For word-initial clusters, children may initially delete a member of a consonant cluster (one-element realization), then preserve the members while producing one in a non-adult manner (two-element realization), and finally they will produce the consonant cluster correctly (correct realization). Other developmental sequences are possible, particularly for word-final consonant clusters.

9. There is an interrelationship between cluster reduction, cluster simplification, and correct productions of consonant clusters. Initially, most children reduce consonant clusters. Over time, the occurrence of cluster reduction diminishes, whereas the occurrence of cluster simplification increases. Simultaneously, the occurrence of correct productions increases, until eventually production is mastered.

10. Despite there being a typical developmental sequence, the acquisition of consonant clusters is marked by reversals and revisions with considerable individual variation.

Summary

Speech-language pathologists are often faced with decisions regarding the speech status of young children. With a body of evidence spanning 70 years, and recent research in the normal speech development of very young children, there is a need for a readily accessible collation of information regarding normal development. The general trends regarding the development of consonant clusters presented in the current overview can inform the assessment, analysis, and intervention decisions of speech-language pathologists. Decision-making regarding the presence of impairment can be assisted with knowledge of the scope of individual variation and data on normal development from a variety of theoretical approaches.

Acknowledgments

This paper presents a portion of work from a doctoral dissertation conducted by the first author at The University of Sydney. The research was supported by grants from The University of Sydney Research Awards and the Charles Sturt University PhD Write-up Award Scheme.

References


Hodson, B. W. (1989). From articulation to phonology: Remedia-


