Assessment of Speech or Sound Production
Includes Articulation and Phonological Disorders

Wisconsin DPI thanks the Virginia Department of Education (VDOE) for allowing DPI to adapt portions of their resource guide for Wisconsin. Wisconsin DPI would also like to thank the following people for contributing to this work: Brenda Forslund, Lisa Kirby-Mangas, Dawn Merth-Johnson, Paula Millikin.

DEFINITION PI 11.36(5)(a), Wisconsin Administrative Code Speech or Language Impairment means: An impairment of speech or sound production, voice, fluency, or language that adversely affects educational performance or social, emotional, or vocational development.

Speech Sound Disorder (Articulation)
Following consideration of the child’s age, culture, language background, and dialect, the child meets all of the following conditions for a speech sound disorder:

a. The child’s speech sound production is documented to be delayed, as evidenced through at least one observation in a natural environment.

b. The child’s speech sound production is documented to be delayed, as measured by a criterion-referenced assessment, such as a developmental scale or a phonetic inventory, or significant discrepancy in performance from typical on a norm-referenced assessment.

c. The child’s intelligibility is below the expected range and not due to influences of home languages or dialect. Intelligibility ratings as documented by school staff or caregivers indicate an impact across environments.

d. Speech sound production is less than 30% stimulable for incorrect sounds.

Phonological Disorder
Following consideration of the child’s age, culture, language background, or dialect, the child demonstrates the characteristics of a phonological disorder, which include both of the following:

a. The child’s phonological process use is documented to be non-developmental or outside of the expected developmental range, as evidenced through at least one observation in a natural environment.

b. The child’s phonological process use is documented to be non-developmental or outside of the expected developmental range, as evidenced by measurement of
either the presence of one or more phonological processes occurring at least 40%, significant discrepancy in performance from typical on a norm-referenced assessment, or both.

C. The child’s intelligibility is below the expected range and not due to influences of home languages or dialect. Intelligibility ratings as documented by school staff or caregivers indicate an impact across environments.
# Table of Contents

**Key Ideas for Speech Disorders**

**Procedures and Tools Assessing Speech**

**Academic Activities**
- Engage in Discussion with Classroom Teacher(s)
- Review Classroom Artifacts
- Observation Tips and Tools

**Contextualized Measurement**

**SLP Probes**
- Case History and Interviews
- Oral Motor Structure and Function

**Criterion-Referenced Assessment**
- Phonemic Inventories
- Developmental Scales
- Phonological Processes (or Phonological Patterns)

**Speech Intelligibility**
- Intelligibility Sample
- Intelligibility in Context Scale (ICS)
- Percentage of Consonants Correct (PCC)
- PCC Connected Speech Sample
- PCC Imitative Sentence Procedure

**Stimulability**

**Norm-Referenced Assessments**

**Understanding Academic Language and Adverse Effect**

**Summarizing Assessment Data**

**References**

**Links for Additional Resources and Tools**
Key Ideas for Speech Disorders

- “An articulation/phonological impairment is characterized by an inability to use speech sounds that are appropriate for a person’s age and linguistic dialect. Such errors in sound productions may interfere with intelligibility, social communication, and/or academic and vocational achievement” (Virginia Department of Education [VDOE] 2020, p.8).

- “Although articulation and phonology are both terms used when describing speech sound production, they are not interchangeable.
  - Articulation can best be described as the movement of the articulators when producing a sound, while phonology is a component of language that controls the patterning of speech sounds.
  - When describing speech sound production errors in terms of articulation, the assumption is that there is a problem with the movement of the articulators which needs to be corrected on a sound-by-sound basis.
  - When describing speech sound production errors in terms of phonology, the assumption is that there is a problem with the patterning of the sounds and it is connected to the meaning of language. In that case, remediation should focus on changing the patterns of sound production in groups, and emphasizing the impact of the change on meaning” (VDOE 2018, p. 61).

- Ensure the IEP team is conducting a comprehensive special education evaluation which includes obtaining information from all of the following: academic activities (including observation), contextualized measurement, SLP probes, as well as norm-referenced assessments, if appropriate.

- Rule out hearing problems with either a review of past screening or current screening.

- Consider the student’s home languages or dialect spoken and utilize the ASHA Practice Portal Phonemic Inventories and Cultural and Linguistic Information Across Languages. Students who demonstrate differences in production of English due to their home languages or dialect spoken should not be penalized (i.e., considered impaired) for these differences. A student who is bilingual or multilingual would be considered for impairment in these areas if the delay is significant in their home language(s).

- Evaluate the student in their home languages unless it is not feasible to do so.

- Use the most recent articulation developmental norms. Currently these are the Crowe and McLeod norms (2020) which represent a compilation of 15 studies of
18,907 children from the United States. Additional versions of these norms can be found at [Wisconsin Articulation Speech Development Infographic](#), which can also be found in the [Links for Additional Resources and Tools](#) section of this document.

- Some areas of assessment may require additional consideration depending on the age of the student. The following guidelines are in the VDOE 2020 guidance (p. 8):
  - “Ages 3-5: Intelligibility, phonological process usage, and stimulability are usually more important than social and vocational considerations.
  - Ages 6-9: Speech sound production norms and stimulability are the typical focus. Social and academic variables should be given stronger consideration.
  - Ages 9 and up: Stimulability and social and academic/vocational considerations are of high importance for this age group.”
- Any documented delay must impact the student’s education (i.e., academic, social, emotional, vocational) in order to identify a student with a Speech or Language Impairment in public schools. See [Understanding Academic Language and Adverse Effect](#).
- For additional information see [Evaluation of Speech: Question and Answer document](#).

### Procedures and Tools for Assessing Speech

#### Academic Activities

**Engage in Discussion with Classroom Teacher(s)**

- Classroom teachers provide important information on progress towards age or grade level academic standards and comparison of typical academic and functional classroom expectations to age or grade level peers.
- Data provided from classroom teachers can include a description of the student’s communication skills in natural settings and how those skills affect classroom functioning, including academic performance or social, emotional, or vocational development.
- Classroom teachers can provide data on the effectiveness of successful accommodations and modifications (e.g. instructional supports that address learner variability) as well as the information on the student’s ability to generalize skills in a variety of settings and content areas.
- A [teacher questionnaire](#) may be a way to capture the functional impact of the speech delay that teachers observe in the classroom.
Review Classroom Artifacts

- Are the student’s running records indicative of an impact of speech sound errors or phonological process use on decoding abilities?
- Are the student’s writing samples filled with spelling errors consistent with the speech sound errors or phonological process use?

Observation Tips and Tools

- Observations should take place in the context of daily activities or routines in multiple settings and situations with different peers and adults such as during a time of social interaction (e.g., morning meeting, recess or lunch) and during academic time.
- Observations should focus on the functional aspect of skills rather than isolated discrete skills (e.g., the student is an effective communicator despite demonstrating some articulation errors).
- Data collection during observations should include quantitative data in addition to qualitative information (e.g., description of what the practitioner is observing). For an example Classroom Observation Tool, go to the DPI Speech-Language Impairment website: Assessment Tools for Speech or Language Impairment.

Contextualized Measurement

Review student data.

- Are districtwide and statewide assessment data below the expected range for the student’s age or grade level?
- Are there other possible reasons why the student may not score within the expected range for their age or grade that are due to issues with instruction, curriculum, or environment? For additional information about the ICEL and RIOT frameworks, go to the DPI Comprehensive Special Education Evaluation website.

SLP Probes

Case History and Interviews

- Families or caregivers should be active participants in the evaluation process, including how the student communicates wants and needs, engages with other children, and transitions between home and community. The Wisconsin Statewide Parent-Educator Initiative (WSPEI) offers Resources, including “Snapshot” forms and “Positive Student Profile” to assist family members in active IEP team meeting participation.
• Interview the student whenever possible. The student can provide firsthand information about peer relationships, attitudes toward school, hobbies and interests, strengths and challenges, sensory concerns, and activities outside of school.

• If a student or family speaks a language other than English, it is recommended a translator is used to obtain case history from family to determine the extent of knowledge and use of the home language(s).

• Family or caregiver information should be respected and used as data sources, including information about their student participating in daily routines in the home or community, identifying behaviors at childcare, results of instructional strategies from home or community intervention programs, medical, or clinical based information. Information from the family is especially important when evaluating culturally and linguistically diverse students. Home visits and use of interpreters as needed can aid in the establishment of a relationship with culturally and linguistically diverse families.

• Evaluators should also interview school staff (including the general education teacher and English as a Second Language [ESL] teacher) regarding the student’s language use across settings (Orellana et al. 2019; Roseberry-McKibbin 2021).

**Oral Motor Structure and Function**

An assessment of a student’s oral structure and function should be completed to ensure that an underlying physical structure or motor issue is not interfering with speech production (VDOE 2018). Structural or functional differences alone do not indicate an impairment; these differences must be determined to affect a student’s speech production when considering a Speech or Language Impairment.

**Criterion-Referenced Assessment**

Criterion-referenced assessments “are...tests [and procedures] that measure an individual's performance against a set of predetermined criteria or performance standards (e.g., descriptions of what an individual is expected to know or be able to do at a specific stage of development or level of education)” (ASHA n.d.a.) and may be either standardized or more informal, clinician-developed in nature. These assessments have a more narrow focus of content when compared to norm-referenced assessments and often have a percentage, mastery/non mastery or pass/fail result. A student would be scored as “pass” if a particular skill was mastered and as “fail” if they did not demonstrate mastery of the content.

Criterion-referenced assessments grew out of a need for better assessment methods. Norm-referenced tests were found to be inadequate for determining present levels of
performance and identifying targets for intervention. They also have limited utility when a student is not represented in the normative sample due to their cultural and linguistic background (McCauley 1996). Criterion-referenced assessments include (but are not limited to) utilizing phonetic inventories or developmental scales or norms, intelligibility samples, Percentage of Consonants Correct (McCauley 1996). In addition, SLPs may utilize norm-referenced assessments as criterion-referenced assessments; in this situation, standard scores would not be reported but proficiency of specific skills would be reported.

The following are types of criterion-referenced assessments commonly utilized by speech-language pathologists:

**Phonemic Inventories**

Phonemic inventories are lists of sounds characteristic of a language or dialect (ASHA n.d.b.). They help identify sounds a student has or is missing in their home language or dialect. The phonemic characteristics of a student’s first system influences the phonemic production of English; these differences should not be considered “wrong” or in error.

**Developmental Scales**

Developmental scales include information about typical speech and language development; the student's skills are compared to the expected age of acquisition of skills (see ASHA’s Typical Speech and Language Development resource for parents). Crowe and McLeod 2020 norms can be used to highlight sounds the student has mastered and has not mastered in consideration of a Speech or Language Impairment. See also Wisconsin Articulation Speech Development Infographic.

Noted Exceptions from VDOE (2020) that are also applicable for Wisconsin: “For students producing lateralized sibilants, using norms to determine if therapy is warranted is not best practice because self-correction does not usually occur with lateralization. There is literature to support not using developmental norms to determine when to provide therapy for lateral /s/” (p. 10).

**Phonological Processes (or Phonological Patterns)**

The following section is directly referenced with permission from VDOE (2020).

“When multiple sounds are in error, phonological processes provide a way to examine patterns of sound errors. Phonological processes go beyond individual phonemes to changes that occur regularly for entire classes or groups of sounds. Processes can be divided into three categories:
1. Whole Word/Syllable Processes change the syllable structure of the word by either taking away a sound(s), adding a sound(s), moving a sound, or a combination of these.

2. Substitution Processes substitute one sound for another, changing something in the manner, place or voicing of the sound.

3. Assimilation Processes are also known as harmony processes as one sound changes to become more like (or exactly like) another sound in the word.

Phonological processes simplify the production of speech and can be part of normal development. When processes continue beyond a developmental stage they may impact intelligibility. Some processes have been shown to have a greater relative effect on intelligibility than others. For example, research shows that final consonant deletion and stopping have a greater impact on intelligibility than velar fronting (Klein and Flint 2006).

Processes like unstressed syllable deletion, reduplication, and assimilation often disappear before age three, while cluster simplification, gliding of liquids, vocalization and stopping tend to persist the longest, up to age five and beyond. Only processes that are not developmental and occur in 40 percent or more opportunities should be noted. However, when there is evidence of at least one process that meets the 40 percent criterion, it is important to document any additional processes used more than 15 percent” (p.10). See the Phonological Processes Chart which can also be found in the Links for Additional Resources and Tools section of this document.

Speech Intelligibility

“Intelligibility is a perceptual judgment that is based on how much of the child’s spontaneous speech the listener understands” (ASHA n.d.d.). Determining intelligibility can be meaningful for identifying the severity of a speech impairment as well as for [monitoring] progress (Allison 2020). Allison (2020) argues that intelligibility can be defined as speech accuracy but is hard to distinguish from comprehensibility, which is affected by gestures, context, physical environment. Comprehensibility and functional communication are terms both used in describing how successful students are in making themselves understood in real-life situations despite difficulties with speech accuracy. A speech disorder is significant when it negatively affects conversational speech intelligibility and interferes with functional communication.

Recent research (Hustad et al. 2020; McLeod et al. 2015; McLeod 2020) acknowledges that intelligibility varies by context (e.g., familiarity of listeners, word or discourse level, amount of background noise) and therefore intelligibility ratings must be interpreted with caution given the these variables. It is beneficial to obtain more than one sample in more
than one context and to consider intelligibility within the context of other assessment activities conducted when considering a Speech or Language Impairment.

**Intelligibility Sample**

The following information was adapted from the 2003 Wisconsin DPI Speech and Language Impairments Assessment Technical Assistance Guide. A student’s conversational speech intelligibility can be assessed using a variety of procedures. One of the most reliable and valid assessment procedures for evaluating conversational speech intelligibility requires calculating the percentage of words understood by the listener from a conversational speech sample (Gordon-Brannan and Hodson 2000; Kent, et. al., 1994; Kwiatkowski & Shriberg, 1992; Peña-Brooks & Hedge, 2000).

A trained speech-language pathologist who is not familiar with the student calculates the percentage of words understood while transcribing a student’s audio recorded conversational or play-based speech sample. Some clinicians recommend calculating “percent speech intelligibility” from a 100 word sample (Gordon-Brannan and Hodson 2000) while others suggest a 200 word sample (Weiss et al. 1987). The formula for calculating percent speech intelligibility is provided below for a 100 word sample. This measure often results in a statement such as “Susan’s conversational speech intelligibility was 63 percent, that is, 63 of 100 words were understood by an unfamiliar speech-language pathologist who listened to [an audio] recorded sample of Susan’s speech during conversation.”

\[
\text{% of Intelligible Words} = \frac{\text{# of Intelligible Words}}{\text{Total # of Words}} \times 100
\]

Data from a variety of clinical and research sources (Gordon-Brannan and Hodson 2000; Peña-Brooks & Hedge, 2000; Vihman & Greenlee, 1987) reveals average speech intelligibility and range expectations vary for typically developing 3-5 year olds. “Children above the age of 4 with intelligibility percentages below 66 percent may be ‘at risk’. The children farther along the continuum toward unintelligible speech would be of greatest concern not only for communication success, but also potentially for problems in developing literacy skills” (Gordon-Brannan and Hodson 2000).

“For young students who are highly unintelligible, Gordon-Brannan and Hodson (2000) suggest an alternative measure of intelligibility using imitated sentences. Some advantages of the imitated sentence measure are: (a) suprasegmental features and some syntactic/morphological and contextual cues are available, (b) it takes less time to administer and score than the continuous-speech procedure, and
(c) the child’s intended utterance is known by the examiner” (VDOE 2020, p. 9). The SLP who is conducting the assessment should be aware that children are typically more intelligible in conversation where the context is known than when imitating sentences.

**Intelligibility in Context Scale (ICS)**

The Intelligibility in Context Scale (ICS) is a parent-completed questionnaire developed in 2012 by McLeod, Harrison, and McCormack to determine functional intelligibility (which is affected by speech sound production and contextual factors). The ICS was created to balance out intelligibility determined in a more clinical context during assessment by documenting intelligibility with a variety of listeners across contexts. It also has been used to determine functional intelligibility in each language if a student is bilingual or multilingual. The ICS utilizes a 5-point Likert scale. Scores of 1 correspond with low intelligibility and scores of 5 correspond with high intelligibility.

Research across 14 countries indicates that typically developing 4- to 5-year-old children are always (i.e., “5”) to usually (i.e., “4”) intelligible, even to strangers (McLeod et al. 2015; McLeod, 2020). The ICS is currently available in over 60 languages. “The overall impact of decreased intelligibility (ICS score of 3 or lower) should be determined by the team with consideration of environment” (VDOE 2020, p.3). Download questionnaires for free at the Charles Sturt University Multilingual Children’s Speech page.

**Percentage of Consonants Correct (PCC)**

The following information was adapted from the 2003 Wisconsin DPI Speech and Language Impairments Assessment Technical Assistance Guide. Percentage of Consonants Correct “differs from intelligibility in that it reflects the segmental accuracy of the child’s production and does not take into account the listener’s ability to understand the message being conveyed” (Allison 2020). Either imitative or connected speech samples may be used when calculating PCC.

**PCC Connected Speech Sample**

PCC Connected Speech Sample, which is closely related to percent of conversational speech intelligibility, provides an index of speech impairment severity: mild, mild-moderate, moderate-severe, and severe. Shriberg & Kwiatkowski (1982) developed the procedures for “Percentage of Consonants Correct” and are outlined as follows:
• Spontaneous samples should include at least 90 different words. If the student is so unintelligible that it is impossible to identify this number of different words, then a single word assessment tool may be used for analysis.

• The PCC is calculated by using the following formula:

\[
\frac{\text{Number of Correct Consonants}}{\text{Number of Correct Plus Incorrect Consonants}} \times 100 = \text{PCC}
\]

Example:

50 consonants produced correctly divided by 200 total consonants attempted

\[
\times 100 = 25\% \text{ (PCC score)}
\]

• Determine the Severity Level by using the following scale (Shriberg and Kwiatkowski 1982) to determine the severity of the disorder:

85-100% mild
65-85% mild-moderate
50-65% moderate-severe
<50% severe

**PCC Imitative Sentence Procedure**

In 2004, Johnson, Weston, and Bain found that an “imitative sentence procedure provided PCC scores that compared favorably to those derived from spontaneous speech” (p. 63) and that the imitative sampling procedure was significantly more efficient than spontaneous speech sampling. Wisconsin DPI has adopted a tool based on Johnson, Weston, and Bain (2004) called Fillable Percentage of Consonants Correct (PCC) Sentence Scoring Form for this type of data collection. The following is a recommendation of procedures to follow:

1. Present sentences using a conversational tone without exaggerated prosodic cues (Johnson, Weston, and Bain 2004).
2. Only consonants are scored, not vowels (i.e., only the consonantal /r/ is scored).
3. Mark errors directly on the list of sentences for efficient scoring.
4. Score only the first production of a consonant if a syllable is repeated (e.g., ba-balloon. Score only the first production of /b/).

5. Do not score consonants if a word is unintelligible or only partially intelligible.

6. Errors include substitutions, deletions, distortions, and additions. Voicing errors are only scored for consonants in the initial position of words.

7. If /ng/ is replaced with /n/ at the end of a word, do not score it as an error. Likewise, minor sound changes due to informal speech and/or selection of sounds in unstressed syllables are not scored as errors (e.g., /fider/ for “feed her,” /dono/ for “don’t know”).

8. Dialectal variations are not scored as errors.

Stimulability

Stimulability is “correctly imitating a sound when given auditory and visual cues after a previously incorrect spontaneous production” (Miccio 2002). There is evidence to suggest that if sounds are stimulable, they can be acquired without specially designed instruction (Miccio et al. 1999; Powell et al. 1991; Miccio 2002). For example, the evidence-based practice of the complexity approach focuses on providing intervention to sounds and classes of sounds for which the student is not stimulable (Storkel 2018). The following are guidelines for determining a student’s stimulability:

1. Only sounds that are absent from the student’s phonemic inventory are tested. If 30% or more of the productions for a target sound are correct, the sound is considered stimulable.

2. Sounds are probed at the sound, syllable, word or sentence level. In each probe, the student is asked to repeat a sound in various word positions and with varying vowels. Each attempt is scored as correct or incorrect.

3. If several sounds are found to be absent from the inventory, the probe may be shortened to contain only one vowel context for the target sounds.

“Sounds that are not stimulable for production (0% correct) are... least likely to be acquired in the absence of treatment” (Miccio 2002), and sounds that are at least 30% stimulable will grow with home practice or reminders from general education teachers. WI DPI has adopted the Fillable Miccio Stimulability Probe. Storkel (2018) and Miccio (2002) have each created informal stimulability probes to assist SLPs in easily obtaining stimulability data for sounds students produce in error. The Storkel procedures (2018) follow those outlined for the Miccio probe (2002). Some norm-referenced assessments have stimulability probes as part of the assessment (Glaspey 2012).
Stimulability may also be considered as part of dynamic assessment (Glaspey 2012; Storkel 2018). Termed “speech adaptability”, this process includes capturing the amount of support and models needed from the evaluator and the corresponding response from the student at varying levels of linguistic complexity (i.e., isolation, word, sentence levels; Glaspy 2012).

**Norm-Referenced Assessments**

Norm-referenced assessments are standardized tests designed to compare and rank order students in relation to one another (ASHA n.d.a.; McCauley 1996). A student's performance is reported as a percentile or standard score that is a comparison to the normative sample. Tests are designed to include test items that those with disorders get wrong while those without disorders get right (McCauley 1996), as the purpose is to discriminate between individuals. Norm-referenced assessments are most often broad areas of assessment (e.g., oral language skills). In selecting norm-referenced assessments, evaluators should review the psychometric properties of the test in order to determine whether each are appropriate for particular students and their area(s) of need.

Information about the Limitations of Norm-Referenced Assessments can be found on the DPI Speech-Language Impairment website: Assessment Tools for Speech or Language Impairment once it is published.

In Wisconsin, “significant discrepancy” means performance on a norm-referenced assessment that meets the cutoff score for a speech or language disorder and is significantly below age- or grade-level expectations relative to a normative sample, often reported as a percentile or standard score. Additional information about Significant Discrepancy can be found on the DPI Speech-Language Impairment website: Assessment Tools for Speech or Language Impairment once it is published.

- Significant discrepancy is based on tests' sensitivity, specificity and cut scores.
- If a test's sensitivity and specificity is 80% or higher, consider the cut score for determining disorder on the instrument but also consider:
  - Is the score within the average range (i.e., within 1 standard deviation of the mean)? If so, consider other information in determining an impairment in the schools.
  - Is it at least -1.5 SD from the mean? This score would indicate moderate to severe impact. The evaluator should still look at social, emotional, academic impact and other factors in determining a Speech or Language Impairment.
Understanding Academic Language and Adverse Effect

When conducting assessments and considering additional information to determine if a student has a speech-language impairment, it is a requirement that there is documentation of the impact of the identified speech-language delay.

Academic impact may be documented by considering any of the following:

- Running records, writing samples, districtwide assessments
- Errors in speech showing up in spelling
- Errors in speech affecting accuracy of decoding
- Delayed phonological awareness skills

Social and emotional impact may be documented by considering any of the following:

- Observations, interviews (including with the student), questionnaires
- Student withdrawing from social situations
- Student being teased by peers
- Student withdrawing from large and small group conversations

Summarizing Assessment Data

- Evaluators should consider information from a variety of sources when determining whether a student has a Speech or Language Impairment.

- When determining a student’s academic or functional performance in any area, the team should not rely on a single data point (i.e. one assessment or test score). Triangulating data is a strategy that can be used to compile multiple (at least 3) types of data from different sources.

- Assessment results should include sufficient information so the IEP team can consider the student’s previous rate of academic growth, whether the student is on track to achieve or exceed age or grade-level standards and expectations, any behaviors interfering with the student’s learning or learning of others, the effectiveness of instructional interventions, and any additional information and input provided by the student’s parents.

- Wisconsin has adopted the Rubric for Speech Impairment, which can also be found in the Links for Additional Resources and Tools section of this document.
References


Links for Additional Resources and Tools

The following are additional resources and tools referenced in this document and relevant during assessment of speech sound disorders.

- [Updated Sound Development Chart](#)
- [Wisconsin Articulation Speech Development Infographic](#)
- [ICEL/RIOT Framework](#)
- [Phonological Process Chart](#)
- [Percentage of Consonants Correct Imitative Scoring Form](#)
- [Miccio Stimulability Probe](#)
- [Rubric for Speech Impairment](#)