Comparing speech assessments: The usefulness of the DEAP as compared to the GFTA-2

Ashley Brianne Pitts

Follow this and additional works at: https://digitalcommons.wustl.edu/pacs_capstones

Part of the Medicine and Health Sciences Commons

Recommended Citation

This Thesis is brought to you for free and open access by the Program in Audiology and Communication Sciences at Digital Commons@Becker. It has been accepted for inclusion in Independent Studies and Capstones by an authorized administrator of Digital Commons@Becker. For more information, please contact vanam@wustl.edu.
This study seeks to evaluate whether the DEAP, a new speech assessment that assesses vowels in addition to consonants, is as effective with children who are deaf and hard of hearing as an older, more established speech assessment, the GFTA-2.
Acknowledgments

I would like to thank my advisor, Christine Gustus, for all of her dedication and guidance in helping me to complete this study. I would also like to thank my secondary reader, Dr. Heather Hayes, for giving of her time and effort throughout this project.
# Table of Contents

- Introduction ........................................................................................................ 2
- Literature Review ............................................................................................. 2
- Method .............................................................................................................. 5
- Discussion ........................................................................................................ 12
- Conclusions ...................................................................................................... 15
- References ........................................................................................................ 17
- Appendix- Parent Letter .................................................................................. 19
**Introduction**

When working with children who are deaf and hard of hearing, it is important to continually monitor their growth to make sure they are progressing at a steady pace. There are a variety of ways in which to assess students, both formally and informally. One advantage to formally assessing students or administering standardized tests is the ability to track a student’s progress over a long period of time. Unfortunately, with many assessments currently available, educators can often become overwhelmed with choices. Many times, once an educator finds an assessment he/she deem reliable, he/she will continue to use that assessment for as long as possible. One such reliable assessment is the *Goldman-Fristoe 2 Test of Articulation*, an assessment that focuses on identifying speech errors in words and sentences.

The *Goldman-Fristoe 2 Test of Articulation*, or the GFTA-2, was first created in 1986, and due to its popularity, it was later updated with clearer pictures and restandardized norms in 2000 to provide a more accurate assessment in relation to the child’s speech errors. This test has withstood the test of time; however, there have been several speech assessments that have been created since then. One such assessment is the *Diagnostic Evaluation of Articulation and Phonology*, or the DEAP, which was created in 2006. Something that makes this speech assessment unique is that it assesses vowels in addition to consonants. This study will attempt to determine if this test is a reliable tool to use to evaluate children who are deaf and hard of hearing by comparing it to the GFTA-2 in terms of what is assessed, ease of administration, and the overall standard scores.

**Literature Review**

Hearing loss generally affects most aspects of speech production as well as the normal development of speech (Ling, 1976); therefore, it is important that educators of the deaf
understand where students are in regards to their development of speech skills. Daniel Ling refers to five basic stages of speech acquisition and development: “undifferentiated vocalization; nonsegmental voice patterns varied in duration, intensity, and pitch; a range of distinctly different vowel sounds; simple consonants releasing, modifying, or arresting symbols; and consonant blends” (2002, p. 113). These five stages are “hierarchical and cumulative” (Ling, 2002, p. 113), meaning these stages must occur in the order in which they are presented because each stage builds upon the previous one. It is important to note that “unless the child can vocalize on demand and can produce a wide range of voice patterns and a variety of vowels, we should not seek to initiate or extend his consonant repertoire” (Ling, 2002, p.113).

The above information points to the fact that it is necessary to know where the child is in his development of the speech sounds (including consonants and vowels) before appropriate speech training can occur. This indicates the need for some sort of assessment tool that assesses more than just consonants to serve as a guide when formatting a suitable plan for intervention.

One notable method for assessing speech in students who are deaf and hard of hearing was developed by Daniel Ling over three decades ago; it is known as the Phonetic Level Evaluation, or PLE (Ling, 1976). The PLE evaluates students’ productions of nonsegmental aspects of speech, vowels and diphthongs, simple consonants, and consonant blends through imitation in four hierarchical tasks (Tye-Murray & Kirk, 1993). As of 1988, this instrument was used in over 180 facilities across the United States (Abraham, Stoker, & Allen, 1988).

A wide variety of speech assessments are currently available on the market today. A few examples include the Clinical Assessment of Articulation and Phonology (CAAP) (Secord, Donohue, & Johnson, 2002), Assessment of Sound Awareness and Production (ASAP) (Mattes, 1998), and Photo Articulation Test, Third Edition (PAT-3) (Pendergast, Dickey, Selmar, &
Some of these assessments focus on the articulation aspect of speech sounds (Bleile, 2002) while others focus on assessing speech sound errors within other aspects of the language system (Hoffman and Norris, 2002). Clinicians must decide which assessment most effectively measures the errors of their students. Other notable assessments include the Goldman-Fristoe 2 Test of Articulation (GFTA-2) and the Khan-Lewis Phonological Analysis, Second Edition (KLPA-2). In a study conducted by Skahan and Watson (2007), educators across the country were contacted to discuss which speech assessments were most frequently used. The authors found that 51.8% of the educators claimed they “always” used the GFTA. Although this assessment is a good tool to use to evaluate the speech skills of students who are deaf and hard of hearing, vowel errors are not assessed in the results.

Vowel production errors are common in children who are deaf and hard of hearing (Levitt & Stromberg, 1983; Monsen, 1976, 1978), and therefore, it is important to assess these errors. One assessment that evaluates vowel errors is the Arizona Articulation Proficiency Scale, Third Revision (Arizona-3) (Fudala, 2000). This measure assesses all of the major speech sounds in the English language (asha.org, 1997-2010). Although this is a good example of an assessment used to evaluate vowels, there are still very few options for assessing vowel errors of children who are deaf and hard of hearing. Because of this, a study was conducted in 1991 to provide suggestions for “supplementing tests with additional stimulus words in order to obtain an adequate sample for vowel analysis” (Pollock, 1991). Recently, an assessment has been published that assesses vowel errors in addition to consonant errors: the Diagnostic Evaluation of Articulation and Phonology, or the DEAP (Dodd, Huo, Crosbie, Holm, & Ozanne, 2006)). The DEAP, published in 2006, is a potentially useful tool for educators of deaf children. Some educators of deaf children who use spoken language have questioned whether this assessment
provides accurate information about speech errors when compared to other well-known tests, such as the Goldman-Fristoe 2 Test of Articulation. The goal of the current study is to determine if this test is a reliable tool for educators of the deaf by comparing it to the GFTA-2. Items being compared between the two assessments are the content that is assessed, the ease of administration, and the overall standard scores of the participants.

Method

Participants

This study included 5 deaf students ages 3 years 3 months to 8 years 9 months. All children use spoken language as their primary communication mode and all were students at a private school for the deaf in St. Louis, Missouri. This private school emphasizes the use of listening and spoken language skills. At the time of testing, the children had been students at the school for different periods of time, ranging from three months to three years. All of the students received auditory information through bilateral amplification: four students wore bilateral cochlear implants and one student wore a cochlear implant and a hearing aid. The students ranged in age from 3.0 and 8.11. The students varied in language and speech skills. However, because the results were not compared between the participants, the variance in abilities was not relevant. Table 1 describes the characteristics of each student who participated in the study.
Table 1

Participant Characteristics

<table>
<thead>
<tr>
<th>Information</th>
<th>Student 1</th>
<th>Student 2</th>
<th>Student 3</th>
<th>Student 4</th>
<th>Student 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age as of Test Date</td>
<td>5.8</td>
<td>8.9</td>
<td>4.3</td>
<td>3.3</td>
<td>4.10</td>
</tr>
<tr>
<td>Age First Amplified</td>
<td>2.0</td>
<td>2.0</td>
<td>1.8</td>
<td>.02</td>
<td>.09</td>
</tr>
<tr>
<td>Age Implanted- Right</td>
<td>4.4</td>
<td>3.4</td>
<td>2.4</td>
<td>1.0</td>
<td>1.4</td>
</tr>
<tr>
<td>Age Implanted- Left</td>
<td>4.10</td>
<td>-</td>
<td>2.10</td>
<td>1.0</td>
<td>2.0</td>
</tr>
<tr>
<td>Age Entered Current School</td>
<td>5.0</td>
<td>5.3</td>
<td>4.1</td>
<td>.02</td>
<td>2.3</td>
</tr>
</tbody>
</table>

*Note.* Current School indicates the private school for the deaf.

Procedures

Data was obtained by administering two different speech assessments: the Goldman-Fristoe 2 Test of Articulation (GFTA-2) and the Diagnostic Evaluation of Articulation and Phonology (DEAP). All of the subjects were given both tests on the same day to ensure no significant progress could be made between the testing periods that could skew the data. However, both assessments were given at different periods during the day, allowing the subjects a short break in between testing sessions.

The first assessment given was the Goldman-Fristoe 2 Test of Articulation, or the GFTA-2. There are three separate subtests that make up this assessment: Sounds-in-Words, Sounds-in-Sentences, and Stimulability. For the purpose of this study, only the Sounds-in-Words section was given. In this section, there were a total of 53 targeted words listed on the response form. Of these words, there were 23 consonants assessed in the initial, medial, and final positions while
the remaining 16 blends were assessed in the initial position with a total of 77 sounds and blends being assessed. The targeted words were made up of 28 single syllable words and 25 multisyllabic words.

During the Sounds-in-Words subtest, participants were shown an easel containing different pictures on each page. Each picture was a colorful, life-like drawing depicting the noun or verb that contained the one or two consonants that were being assessed. The participants were prompted to produce the targeted word through the use of stimulus questions. Examples of stimulus questions included, “What is this?” or “What do you call this?” The targeted responses were listed on the response form. If the participants gave a response that did not match the target, the examiner replied with either a prompt- (“Tell me another name for it”) or a cue. During a cue, the examiner gives the targeted word at the beginning of a two sentence phrase. This allows the student to hear the word but not focus on how the examiner is articulating it because he is listening to the remaining information. An example of this would be, “These are pajamas. You wear them at night when you go to sleep. What are these?” If the student was still unable to produce the targeted word after the intervention strategies, the examiner modeled the appropriate word aloud for him. When this occurred, the examiner made a note on the response form indicating the word was produced in imitation.

The second assessment given to the subjects was the Diagnostic Evaluation of Articulation and Phonology, or the DEAP. There are two separate response forms for this assessment: the Articulation Assessment Record Form and the Phonology Record Form. This study focused solely on articulation; therefore, the phonology portion of the assessment was not given. The articulation portion of the assessment was broken into three subtests: Articulation Single-Word Production, Phoneme Stimulability, and Oral Motor Screen. For the purposes of
this study, only the Articulation Single-Word Production subtest was used. The targeted responses for this subtest consisted of 24 consonants assessed in the initial and final positions and 2 blends assessed in the initial position of words. This test also assessed 16 vowels in the medial position, with a total of 67 sounds assessed. This portion primarily assessed words following the consonant-vowel-consonant pattern; however, of the thirty targeted words, there were four multisyllabic words included.

During this portion, the participants were shown an easel with a different picture on each page. The pictures depicted were bright and colorful and appeared hand-drawn in a child-like fashion. At the beginning of the assessment, the participants were prompted with, “We’re going to look at some pictures, and I want you to tell me what they are.” The students continued to name each new picture as the pages were turned while the examiner noted the errors on the response form. If the student did not know the correct word for the target, the examiner provided cues similar to those given during the Goldman-Fristoe 2 Test of Articulation. Once again, if the student had to imitate the targeted word, the examiner made note of it on the record form.

Results

The standard scores for each assessment were factored according to the assessment protocols. First, an error score was obtained and then translated into a standard score by use of charts provided by each assessment guide. A mean standard score of 100 (standard deviation=15) indicates average performance for hearing children.
Table 2 shows the standardized scores of the five students on the DEAP and GFTA-2. It is clear that the students’ performances were quite similar across both assessments. The scatterplot in Figure 1 summarizes the positive correlation between the two test occasions, \( r = .97, \ p < .01 \).
When looking at the assessment results, it is important to keep in mind that a standard score of 100 is the mean and not a perfect score. A standard score of 100 only indicates the child has achieved average performance as compared to other students of the same age. When taking into account the standard deviation of 15 on each assessment, all of the students except for one fell 1 standard deviation or more below the mean on both standardized assessments.

Because the DEAP assesses vowels in addition to consonants, it was important to factor two separate percentages for these two segments of the test. Below is the information gleaned when two separate percentages are factored for each student. PCC is an abbreviation for Percentage of Consonants Correct, and PVC stands for Percentage of Vowels Correct.

Table 3

<table>
<thead>
<tr>
<th>Percentages</th>
<th>Student 1</th>
<th>Student 2</th>
<th>Student 3</th>
<th>Student 4</th>
<th>Student 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCC</td>
<td>63</td>
<td>82</td>
<td>43</td>
<td>69</td>
<td>84</td>
</tr>
<tr>
<td>PVC</td>
<td>94</td>
<td>100</td>
<td>94</td>
<td>97</td>
<td>97</td>
</tr>
</tbody>
</table>

Students made few to no vowel errors. The vowel assessment portion of this test was not a large factor in the overall standard scores as the students were at ceiling performance. However, it is important to note that many children with hearing loss continue to make vowel errors, even if this group tested did not.
Another factor to consider when analyzing the overall standard scores between the two tests is the percentage of consonants correct in each assessment. Table 4 details the percentage of consonants correct in both the DEAP and the GFTA-2.

<table>
<thead>
<tr>
<th>Assessments</th>
<th>Student 1</th>
<th>Student 2</th>
<th>Student 3</th>
<th>Student 4</th>
<th>Student 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEAP</td>
<td>63</td>
<td>82</td>
<td>43</td>
<td>69</td>
<td>84</td>
</tr>
<tr>
<td>GFTA-2</td>
<td>48</td>
<td>78</td>
<td>32</td>
<td>51</td>
<td>71</td>
</tr>
</tbody>
</table>

The information gleaned from Table 4 leaves one to question why there is such a large difference in the percentages correct for some of the students on each test. Upon further investigation, it was determined the reason for this disparity was due to the blends that were assessed with the GFTA-2. There were only two blends assessed in the DEAP; however, there were sixteen blends assessed in the GFTA-2. It was found that four out of the five had significant difficulties with the blends. Student 5 only missed seven of the sixteen listed blends. The new results were then calculated by removing the blends and recalculating the percentages. Table 5 indicates the final results of the percentage of consonants correct in both the DEAP and the GFTA-2 (percentages given for scores calculated with and without the blend errors).
Table 5

*Comparing Consonant Percentages*

<table>
<thead>
<tr>
<th>Assessments</th>
<th>Student 1</th>
<th>Student 2</th>
<th>Student 3</th>
<th>Student 4</th>
<th>Student 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEAP</td>
<td>63</td>
<td>82</td>
<td>43</td>
<td>69</td>
<td>84</td>
</tr>
<tr>
<td>GFTA-2 with blends</td>
<td>48</td>
<td>78</td>
<td>32</td>
<td>51</td>
<td>71</td>
</tr>
<tr>
<td>GFTA-2 without blends</td>
<td>66</td>
<td>87</td>
<td>53</td>
<td>68</td>
<td>86</td>
</tr>
</tbody>
</table>

The results of Table 5 allows one to once again see how similar the performance is when the blends are removed from the equation. This table helps the reader to understand that the original factors were not incorrect; there was only a gap due to the large number of blends assessed in the GFTA-2.

**Discussion**

The results from this study suggest that, the DEAP is a reasonable assessment for children who are deaf and hard of hearing, as it produces consonant production results that are similar to those of the GFTA-2, but includes information about vowels, which is lacking in the GFTA-2. This allows an educator to administer either test with confidence, knowing that whichever test they use will produce similar results when compared to the other assessment. Educators of oral deaf children have another choice when it comes to choosing an appropriate speech assessment for their students.

Knowing how much the blends affected the score helps one to realize the importance of using an assessment that measures a student’s ability to articulate all of these sounds. This information makes it seem as though the GFTA-2 would be the best possible choice to use with
students because it assesses a wide variety of blends while the DEAP only assesses two blends. However, the GFTA-2 does not assess vowels like the DEAP does; therefore, it appears as though both tests have both a strong and weak component to them. Fortunately, with the high correlation factor, educators can choose which assessment to give. If they have a student who is exhibiting various vowel errors, then perhaps the DEAP would be a more appropriate assessment to give. If educators have an older student who is working on blends or struggling with different blends, the GFTA-2 would seem like the better choice. Overall, educators can be certain that whichever test they choose to give will produce similar results when compared to the other speech assessment.

Nearly half of the GFTA-2 is composed of multisyllabic words whereas the percent of multisyllabic words on the DEAP is fewer in number. Instead, the DEAP is mostly composed of words following a consonant-vowel-consonant pattern. The large number of multisyllabic words represented on the GFTA-2 can prove to be a challenge when assessing students with a limited vocabulary. In fact, when giving these two tests, the examiner noted when a word had to be produced in imitation. It was determined that, on average, two more words had to be imitated by the student on the GFTA-2 than on the DEAP, due to the fact they were unsure of the vocabulary word’s meaning. Although this certainly does not seem like a high number, it is important to remember that the goal of a speech assessment is for a child to produce the word without any imitation so that the results are valid and can be used to determine the most appropriate intervention strategies. Therefore, even the smallest number of words produced in imitation can have an effect on the overall score because it is not giving the examiner a true sense of the students’ abilities.
Another factor to consider when comparing these two assessments is the different types of pictures each test provides. Throughout the test, the students are expected to look at the easel and label the picture or pictures they see on the page. Each assessment has its own set of pictures. The DEAP utilizes a short easel that contains only one picture per page. The images appear to be almost hand-drawn, child-like in a way whereas the GFTA-2 utilizes a large easel that contains either a single picture or a small picture scene on each page that the students use to answer several questions while the examiner points to each individual item. Both assessments have full color photos to hold the interest of the students while testing. In theory, the idea of using one picture for several words is wonderful; however, with the full scene in front of them, it might be easy for students to get caught up in the picture, rather than focus on the task at hand. It is important to remember those children who can sometimes be over-stimulated by too much visual information; the picture scenes might be too much for them to handle.

The knowledge of all of the shortcomings listed here in regards to these two assessments leads one to question which assessment would be the most appropriate to give due to the fact that neither seems to be the perfect assessment. One option that could perhaps be the best course of action would be to give these two assessments in conjunction with one another. Of course the two tests have a high correlation rating when compared to one another, allowing for an examiner to give either test knowing that the results on the consonants portion of the tests will be similar regardless of which test is given. However, by giving both tests in conjunction with one another, the examiner is provided with the most information about the student, allowing for the best overall speech assessment.
Conclusions

Both the Goldman-Fristoe 2 Test of Articulation and the Diagnostic Evaluation of Articulation and Phonology are invaluable tools when working with students who are deaf and hard of hearing. The findings from this study conclude that the articulation portion of the two tests are positively correlated, allowing educators to feel comfortable giving either test with the knowledge that the results are accurate. However, because data from a single speech assessment is not sufficient for deciding on future goals (Eisenberg & Hitchcock, 2009), these two tests should be used in conjunction with one another. This will provide the examiner with a broader range of information on students’ speech capabilities and shortcomings.

Currently these assessments serve as important resources; however, it is important to note that the field of deaf education is constantly evolving. Therefore, the assessments that are used to evaluate students who are deaf and hard of hearing should be continually updated to reflect the current needs of the population being served. Although this study sought to evaluate another speech assessment to use when evaluating children who are deaf and hard of hearing, the overall sample size was very small. In order to better evaluate new assessments, future studies should seek to utilize a broader sample size to provide educators with the information needed to make an appropriate decision regarding these assessments.

In the future, the author would like to see a variety of follow-up studies as a result from the data collected. One study would consist of this study being conducted again with a larger sample size that is representative of various educational placements available for students who are deaf and hard of hearing, such as private oral schools, self-contained classrooms in public schools, and children who are mainstreamed into general education classrooms. Another study would be to focus on the vocabulary contained within both tests to see whether or not the DEAP
does contain easier vocabulary for younger children. The examiner could focus on the amount of words produced in imitation on either test and take into account whether the amount of words produced in imitation has any effect on the standard scores obtained on both assessments. A final thought would be to assess the phonology portion of the DEAP in comparison to the phonology assessment used in conjunction with the GFTA-2: the Khan-Lewis Phonological Analysis (KLPA-2). This would help educators of the deaf to learn whether or not the phonological portion of the DEAP has a positive correlation with the KLPA-2, allowing for another assessment to be used with confidence when assessing phonology in children who are deaf and hard of hearing.
References


DC: AG Bell Association.


Appendix

November 30, 2009

Dear Moog Parents,

I am currently pursuing my Masters of Science in Deaf Education at Washington University. As part of my graduation requirement, I must complete an independent study. As part of the Independent study, I am required to choose a topic and conduct research that is relevant to the field of deaf education. I have chosen to research how reliable two speech assessments are when compared to each other. The first speech assessment is the Goldman-Fristoe 2 Test of Articulation (Goldman-Fristoe). It is a well-known and reliable test used in many schools, including the Moog Center. The second speech assessment is the Diagnostic Evaluation of Articulation and Phonology (DEAP). This is a newer test that assesses many of the same things as the Goldman-Fristoe. The major difference is that the DEAP also tests vowels. I want to find out if the DEAP is as good or as reliable as the Goldman-Fristoe so that educators can give the DEAP to students with confidence that it assesses everything the Goldman-Fristoe does and more! The information I learn from this study will be presented in May at a workshop in front of my peers and professors.

For my study, I need to give both of these tests to students in a variety of age levels (between the ages of three to eight). The reason your child is being asked to participate in this study is because your child is between the ages of three and eight and attends the Moog Center. I will give the Goldman-Fristoe to your child on any given day, and then within two weeks, give the DEAP to that same child. It is important to give these tests very close together so that they do not learn so much new information that it might skew the data. Both tests will take about thirty minutes to complete. The students will be asked before every testing session if they wish to participate. They are allowed to say no. The student can also stop at any point throughout the test if he/she wishes to. The student will receive a prize whether he/she completes any part of the test that day or not. Please remember, this testing time is meant to be fun as well as productive. I want your child to enjoy this experience.

Your child’s name will not appear on any documents that will be seen by anyone other than myself and Chris Gustus, my project supervisor. All of the information I gather and present at the workshop will be referenced in graphs and tables. If I do mention individual children’s scores, they will only be referred to as Child A, Child B, etc. Your child’s identity will not be known by anyone other than myself or Chris Gustus.

If you have any questions, you may contact me at pittsas@wusm.wustl.edu. You may also contact my project supervisor, Chris Gustus, at 314-692-7172 or cgustus@moogcenter.org.

Thank you for your support,

Ashley Pitts
Graduate Student, Washington University

[Signature]

HRPO
Approved: 2/4/10
Please sign and return to your child's teacher

I, as the child's parent or legal guardian, have carefully read the information regarding the research study. I understand my child, ________________________, will only be tested at The Moog Center for Deaf Education, and he/she will only be given the Goldman-Fristoe 2 Test of Articulation and the Diagnostic Evaluation of Articulation and Phonology. My child will also be given a personal choice as to whether or not he/she wants to participate that day. My child will not be punished for choosing to not take part in the study. I understand my child's identity will be kept private throughout the research study process and during the May workshop; the only two people with access to my child's information are the tester, Ashley Pitts, and her project supervisor, Chris Gustus. I know I can contact either person at any point with questions, comments, or concerns regarding the study.

_____ I give permission for my child to take part in this study.

_____ I do not give permission for my child to take part in this study.

________________________________________
Parent/Legal Guardian