Analysis of 2007 STAMP Results:
New Jersey Grade Eight Proficiency Assessment in World Languages

Executive Summary

The New Jersey State Board of Education has mandated that districts provide students with world languages instruction at the K-8 level since 1996. However, the implementation of this ambitious policy is threatened by the lack of a reliable and valid system of accountability that measures student achievement and the development of quality programs.

In an effort to examine the impact of a valid accountability system on elementary world languages programs, the New Jersey Department of Education (NJDOE) sought and received a FLAP grant from the U.S. Department of Education to provide proficiency testing for Grade 8 students. The results from statewide testing can be used to drive program improvement by assisting districts in determining whether their current curricular design will enable Grade 8 students to demonstrate proficiency at the ACTFL *Novice-High* level, the level targeted by the NJDOE for demonstration of language proficiency at the end of Grade 8 and to meet the current high school graduation requirement.

This report covers Year 2 of the grant period and builds on studies performed in Year 1. It includes a greater number of participating schools and districts, an additional language tested and more detailed information about the each student's language program.

Year 2 proficiency testing of New Jersey Grade 8 world language students was conducted using the Standards-based Measurement of Proficiency (STAMP) developed by the Center for Applied Second Language Studies (CASLS) at the University of Oregon and administered by Language Learning Solutions (LLS). Testing took place between April 2 and June 15, 2007. A total of 21,786 Grade 8 students were tested in 166 schools and 108 districts. The languages tested were: French, German, Japanese, Italian and Spanish. Statewide, 11.9% of the 21,786 students who took the test met the state standard for reading and 25.2% for speaking.

Further analysis of this data revealed six key findings which can be used to improve NJ world language programs:

1. Language programs must provide sufficient hours of instruction in order for students to meet the *Novice-High* proficiency level at the end of Grade 8.
2. Language programs which meet several times each week during the whole school year are generally more effective than an equivalent number of hours in a partial year program.
3. Speaking proficiency is much higher than reading proficiency.
4. Students attending schools in all District Factor Groups have similar speaking proficiency after 5 years of instruction. However, reading proficiency lags for students in lower DFG schools.
5. Heritage students are more likely to meet state standards.
6. Heritage students show similar levels of proficiency across all DFG categories of schools.

Each of these key points is discussed in depth below.

Key Point #1:
Language programs must provide sufficient hours of instruction in order for students to meet Grade 8 state standards.

Background:
For most subjects taught in Grade 8, such as math, language arts and science, students in all districts begin study before Grade 6. This is not the case with world languages programs despite a state requirement to offer world languages in Grades K-8. Currently, districts choose from a variety of program models beginning at different grade levels. Consequently, the total hours of language instruction received by Grade 8 students varies significantly from district to district.

Purpose of analysis:
The 2006 study of NJ students found that Grade 8 scores improved significantly when language instruction began before Grade 6. In 2007, teachers submitted more detailed information about the number of hours of language instruction which their students had received that continues to support the 2006 finding.

Design of analysis:
This study considers only non-heritage learners. For most of these students, the language classroom is their only source of exposure to the language. The 2007 data was detailed enough to estimate each student's lifetime hours of language instruction. This information was used to determine the number of 'instructional units' (IUs) of language instruction. An instructional unit is 108 hours of instruction, i.e., the number of hours received in a program meeting 3 hours per week for 36 weeks per year. By measuring total instructional time in IUs, different program models can be compared easily.

The chart shows that increasing the hours of instruction significantly increases the number of students meeting the state standard. Specifically, students need at least 5 IU of instruction (540 hours) before high rates of success are reached. (The difference between speaking and reading proficiency is discussed in Key Point #3.)
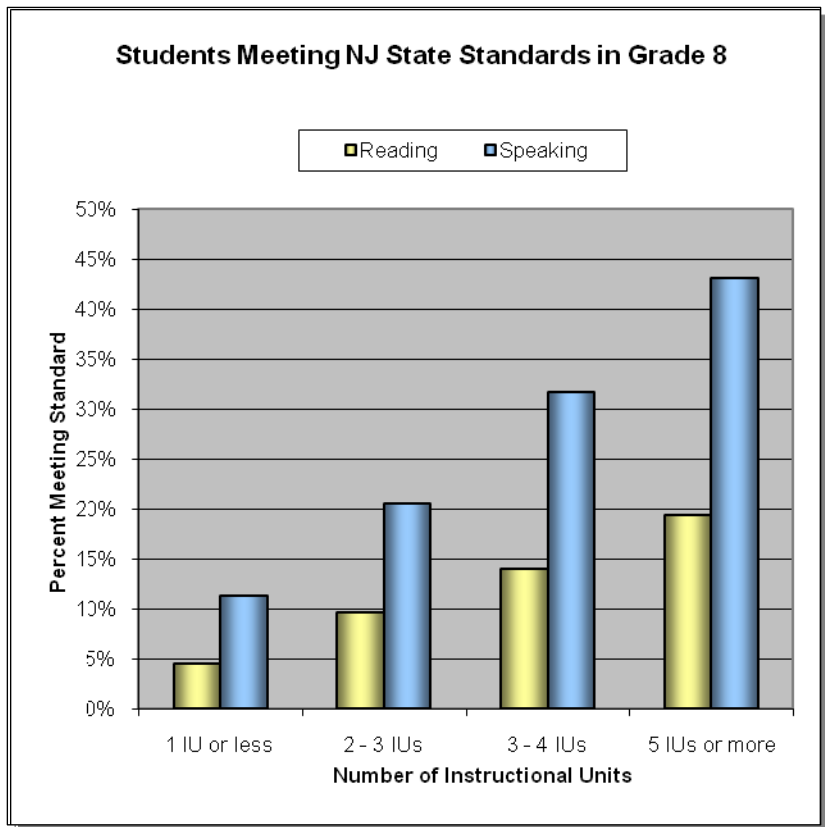
Conclusion:
These results emphasize the need for sufficient instructional time for language study during Grades K-8 if students are to succeed in meeting Grade 8 state standards. They highlight the importance of beginning language instruction as early as possible, ideally, prior to Grade 6. Students who begin instruction at Grade 6 or later will need to devote a significant portion of their classroom hours to language study in order to meet the standard. If this is not possible for practical reasons, language instruction should begin prior to Grade 6, as with other subjects.

Key Point #2:

Language programs which meet several times each week during the whole school year are generally more effective than an equivalent number of hours in a partial year program.

Background:
New Jersey Grade 8 world language students are enrolled in a variety of program models throughout the state. These programs vary by number of weeks per year, number of sessions per week and number of minutes per session.

Purpose of analysis:
This study compares full and partial year program models.

The 2006 study focused on how the amount of time committed to instruction in a student's current language program model affects proficiency. Specifically, we examined the effect of the number of sessions per week and the number of weeks per year. The more effective programs were those which ran more than 21 weeks per year and met at least 3 times per week. **Overall, for a program to lead to significant student proficiency outcomes, there needs to be a minimum of 5,000 minutes of instruction time spread across the school year.**

This analysis is intended to confirm the 2006 results using the more detailed student information available in 2007.

Design of analysis:
Only non-heritage learners are included. Students who have more than 324 lifetime hours of language instruction (or more than 3 instructional units of 108 hours) are excluded. Students in partial year programs rarely reach high numbers of total hours of instruction. This allows a more equitable comparison of partial year and full year program models. Reading and speaking skills were considered separately.

Partial year programs are those running 20 weeks a year or less; programs running 21 weeks a year or more are considered full year programs. Programs were also divided into those

Students Meeting NJ Standards in Grade 8 for Speaking by Program Model

meeting 2 or fewer times per week and those meeting 3 or more times per week. The percent of students in these programs who meet the NJ state standard for speaking proficiency are shown in the chart. Reading and speaking results are similar and shown in the table below.
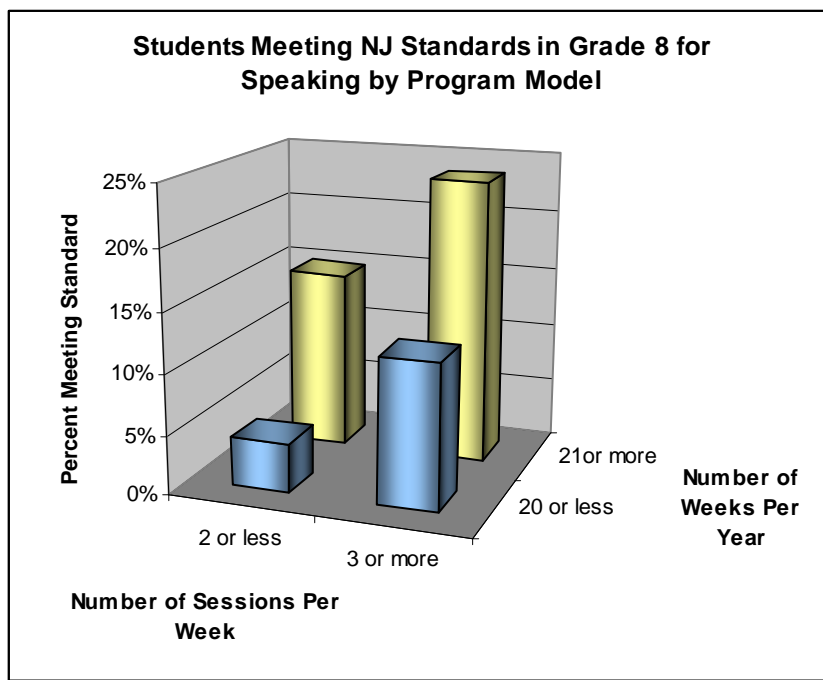
Percent of Non-heritage Students Qualifying at Benchmark Level 3			
Program Model		**Skill**	
Weeks per Year	**Sessions per Week**	**Reading**	**Speaking**
20 or less	2 or less	4.1%	4.1%
21 or more		8.9%	14.9%
20 or less	3 or more	6.2%	12.1%
21 or more		11.1%	23.7%

Conclusion:

These results show the need for sufficient intensity of language instruction if students are to meet the state standard in Grade 8. They highlight the need for consistent instruction throughout the entire school year. **Even when the total number of hours per year is held constant, students perform better when instructional time is evenly distributed across the school year.** This distribution of time requirement is in addition to the need to devote a minimum number of hours to language instruction discussed in Key Point #1.

Key Point #3:
Speaking proficiency is much higher than reading proficiency.

Background:
As in 2006, New Jersey Grade 8 students were tested in both reading and speaking skills using the STAMP assessment. The reading test consisted of a series of multiple-choice items delivered using a computer-adaptive algorithm. At the end of the reading test, students provided speaking samples in response to a standard set of prompts. These were scored using trained human raters.

The students' reading and speaking scores were reported separately as 'Benchmark Level Qualified.' For example, a student might receive a speaking score of 3 and a reading score of 2, indicating that he or she qualified at Benchmark Level 3 in speaking and Benchmark Level 2 in reading. These Benchmark levels are based on the ACTFL Guidelines, as are the NJ state standards. In this example, qualifying at Benchmark Level 3 or greater indicates that the student has met the New Jersey state standard of ACTFL level *Novice-High* or better for speaking. Qualifying at Benchmark Level 2 indicates that the student's reading proficiency is ACTFL *Novice-Mid*. This student has not met the NJ state standard for reading.

Purpose of analysis:
Reading and speaking are distinct language skills which were separately tested and scored. This analysis compares the percentage of students meeting the state standard for each skill.

Design of analysis:
All NJ students who took STAMP in 2007 are included. The chart below shows the counts and percentages of students qualifying at each Benchmark Level for each skill.

Number of Students Statewide Qualifying at Each Benchmark Level

Skill		Benchmark Level						
		NE	0	1	2	3	4	5
Reading	Count	313	-	10,617	8,274	2,258	249	75
	Percent	1.4%		48.7%	38.0%	10.4%	1.1%	0.3%
Speaking	Count	-	3,025	4,990	8,290	5,070	388	23
	Percent		13.9%	22.9%	38.1%	23.3%	1.8%	0.1%

Note: 'NE' indicates that too few reading items were completed to establish a benchmark level, while '0' indicates that too few completed speaking responses were submitted to determine a benchmark level. Responses to two speaking item prompts were needed for grading purposes.

Statewide, 11.9% of the 21,786 students who took the test met the state standard for reading; 25.2% met the standard for speaking. The higher level of speaking proficiency is statistically significant. However, these results are not statistically significantly different from the 2006 results.

The levels of proficiency for both reading and speaking were also statistically different for different languages. However, these results should be interpreted with caution. The number of students studying Japanese was extremely low. Differences between German, French, Italian, and Spanish learners might be accounted for by differences in curricula, programs and districts choosing to offer the language.

Conclusion:
As in 2006, more than twice as many students meet the standard for speaking proficiency as compared to reading proficiency. The increased number of students tested led to slightly lower percentages in both skills as compared to 2006, but the differences are not significant.

Key Point #4:
Students attending schools in all District Factor Groups have similar speaking proficiency after 5 years of instruction. However, reading proficiency lags for students in lower DFG schools.

Background:
The 2006 study showed that students' proficiency scores depend greatly on which school they attend. To better understand these differences, information about each school's District Factor Group is included in the present study and is compared with the number of hours of instruction.

Purpose of analysis:
The goal of this analysis is to determine how much variation in student performance is accounted for by the program model alone and how much by factors particular to the student's school.
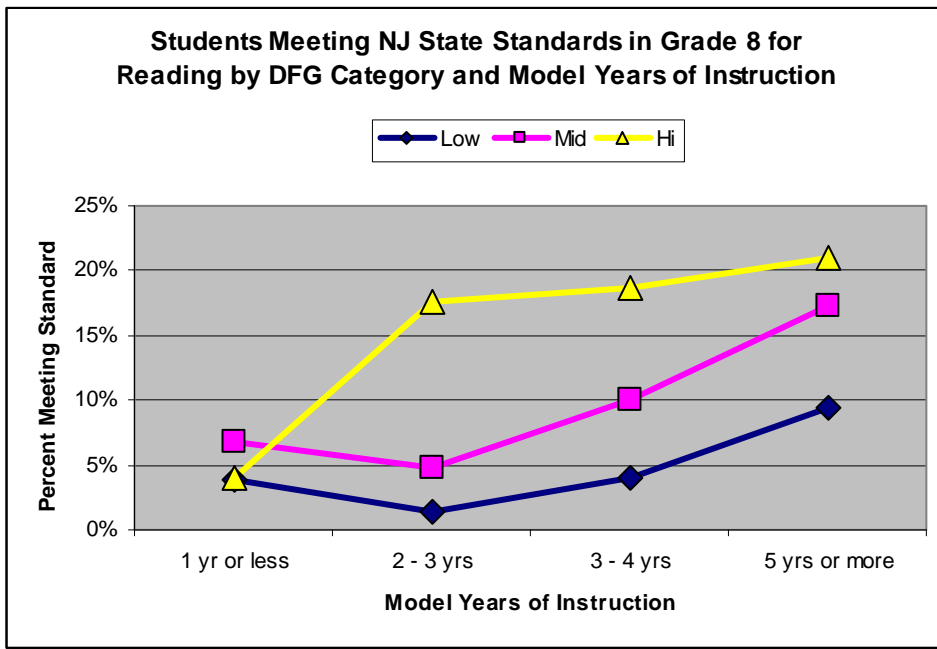
Design of analysis:
The study examines only non-heritage students, as these students must depend on their schools in order to acquire world language skills.

Schools were categorized as belonging to a 'Low', 'Mid', or 'High' District Factor Group, as follows: Low includes DFG A, B, CD, and DE; Mid includes FG and GH; and High includes I and J. Students were then categorized by their total lifetime hours of instruction using an 'instructional unit' of 108 hours (IU), as discussed above. Reading and speaking skills were analyzed separately.

The first chart shows the results for speaking. Initially, Low and Mid DFG schools lag the High DGF schools. However, after 5 IUs of instruction, all groups are statistically indistinguishable. This suggests that by the time many students have acquired solid speaking skills, the effects of attending a particular school disappears. This contrasts with reading proficiency shown in the second chart.
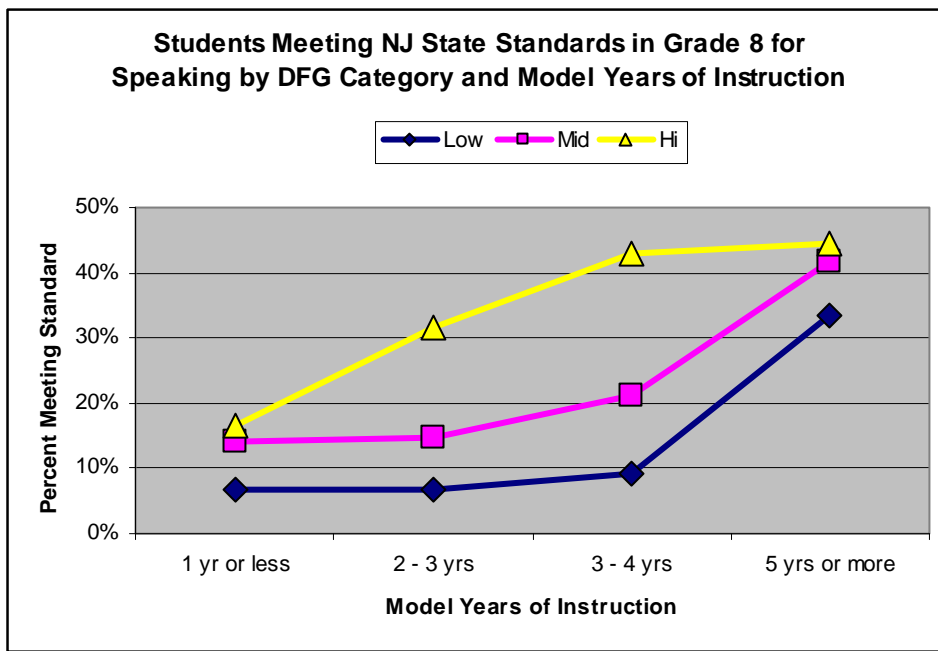
The chart of reading proficiency shows that students in Low DGF schools never catch up with their peers in High DGF schools. Students in Mid DGF schools start out slower than those in High DGF schools, but catch up within 5 instructional units of instruction time.

Students Meeting NJ State Standards in Grade 8 for Reading by DFG Category and Model Years of Instruction

Although students in Low DGF schools acquire equivalent skills in speaking, given the same amount of instruction time, most students in these schools never receive sufficient instruction to acquire high levels of skill. **The chart below shows that Grade 8 students in Low DGF schools have received, <u>on average</u>, only one third the number of hours of instruction as students in High DGF schools.**

Median Lifetime Hours of Foreign Language Instruction of Grade 8 Students by DFG Groupings
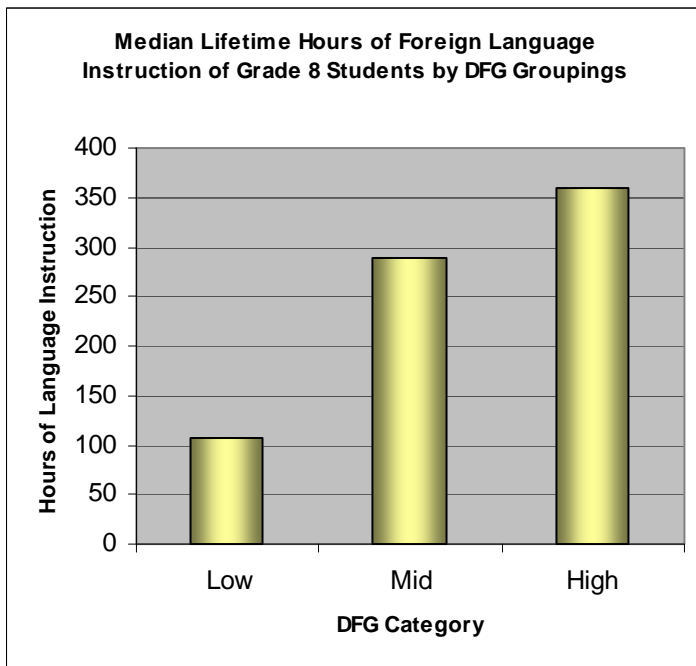
<u>Conclusion</u>:
Given equal amounts of instruction, students in all schools can acquire the same levels of speaking proficiency. Reading skills lag in lower DGF schools. This suggests that reading skills pattern with other academic subjects, while high levels of speaking proficiency are achievable with a sufficient amount of instruction time.

Key Point #5:
Heritage students are more likely to meet state standards.

Background:
Students come to their Grade 8 language classrooms with widely differing world languages experiences. One such experience is exposure to, or interest in the language because it is the heritage language of the student's family.

Purpose of analysis:
For heritage students, much exposure to the language lies outside the students' school experiences, yet it is likely to have a strong effect on proficiency scores. This analysis attempts to quantify the amount of this effect. It builds on the analysis of 2006 to include heritage students for all languages.

Design of analysis:
The 2006 study examined only Spanish heritage students studying Spanish. The present study includes Italian heritage students studying Italian, French heritage students studying French, and German heritage students studying German.

Heritage learners are statistically distinct from non-heritage learners, as shown in the chart. Heritage learners have the highest proficiencies in both reading and speaking. For speaking, 27.1% met the state standard, and this was significantly higher than reading, where 16.2% met the standard. Non-heritage students scored lower; 10.9% met the state standard for reading and 24.7% for speaking.
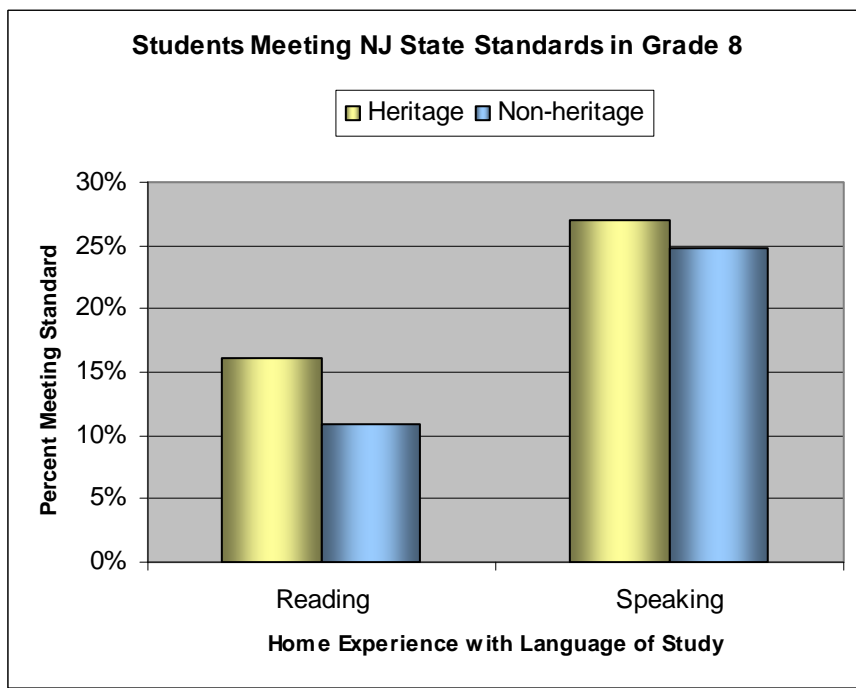
Conclusion:
As in 2006, students studying the heritage language of their families are much more likely to meet state standards both in reading and speaking. However, the effect is not as strong as in the previous study, which included only Spanish learners. This suggests that Spanish heritage students have more exposure to Spanish outside the classroom than heritage students of other languages. This greater exposure was reflected in greater success in meeting state standards.

<u>**Key Point #6**</u>:
Heritage students show similar levels of proficiency across all DFG categories of schools.

<u>Background</u>:
As discussed in Key Point #4, non-heritage students in schools in higher DFG categories show higher levels of proficiencies in most programs. Since heritage students are exposed to the language outside the classroom, this may not be true for these students.

<u>Purpose of analysis</u>:
This study examines whether the number of heritage students meeting state standards is affected by the DFG of the school they attend.

<u>Design of analysis</u>:
Students' schools were categorized as belonging to a 'Low', 'Mid', or 'High' District Factor Group, as discussed in Key Point #4. Only results for speaking are reported here, although a study of reading proficiency showed similar results.

As shown in the chart, both heritage and non-heritage students show similarly high levels of speaking proficiency when they attend higher DFG schools: 39.3% of heritage students meet state standard, 37.0% of non-heritage students. However, the number of non-heritage students meeting standards falls off sharply at mid and low DFG schools: 20.3% for Mid, 8.4% for Low. This is not true for heritage students. Heritage students maintain fairly high levels of success across both Mid and Low DFG schools: 26.2% for Mid, 25.3% for Low.
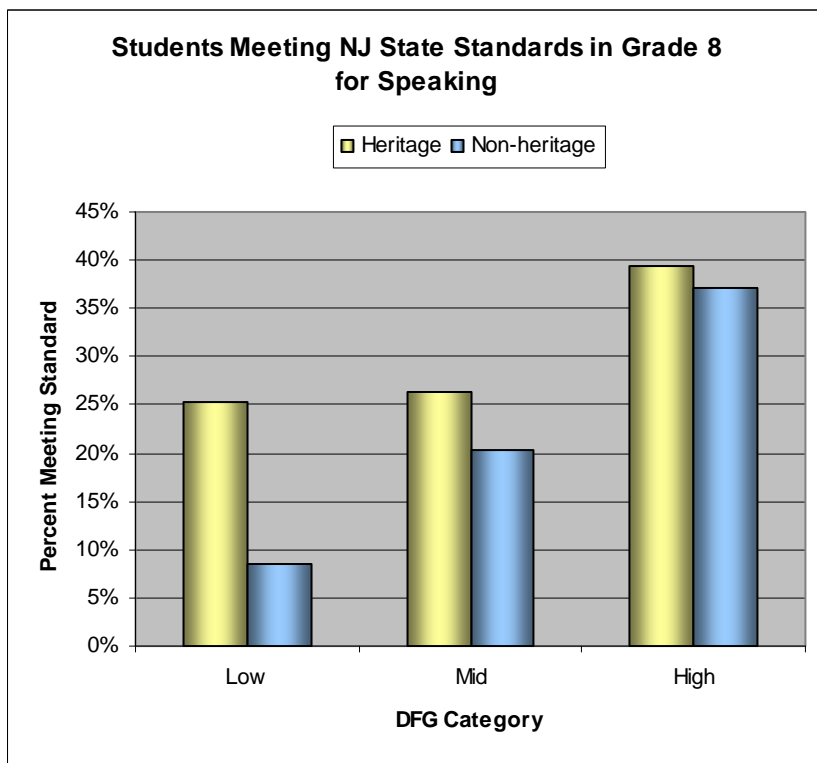
<u>Conclusion</u>:
Heritage students are more likely to meet state standards, irrespective on the DGF category of the school they attend. This suggests that the exposure to the language which they receive outside the classroom helps them meet state standards.