

CASE STUDY : Netherlands Ministry of Justice

Testing Language for Immigration Requirements

Need: Dutch Parliament Mandates New Language Proficiency Requirement

After a prolonged political debate, the Dutch parliament passed a policy in November of 2004 requiring that all persons seeking to immigrate to The Netherlands be able to demonstrate proficiency with spoken Dutch. The Ministry of Justice, which was responsible for implementing this law, turned for guidance to its official partner, CINOP, the Dutch institute for innovation of education and training.

Challenge: Finding a Practical Automated Solution that is as Good as Human Raters

The Ministry required a test for simple spoken Dutch that could be implemented in a consistent and fair manner at 110 embassies and 27 consulates around the globe. However, it would be logistically very difficult to recruit and train expert linguists to administer standardized interview tests in all of these locations. The Ministry was also concerned that it would be difficult to ensure reliability and consistency across so many raters and regions. Presented with these challenges, the Ministry determined it needed a more practical standardized test that could be automatically scored in a valid and reliable way.

To accomplish this highly specialized task, CINOP created a consortium of language testing experts, which included Language Testing Services (LTS), a Netherlands-based company with world-class experience in assessment, and Pearson, a leading expert in automated spoken language testing.

“We needed a test that was extremely accurate but simple to deploy,” said Petra Poelmans, PhD, Consultant with CINOP. “We chose Pearson because it is the only company that could meet these high standards and expectations.”



Solution: Implement a Practical and Valid Test for Dutch Speaking Skills

Leveraging the test development expertise and patented technology used to create its automated tests of spoken English and Spanish, Pearson developed the Toets Gesproken Nederlands (TGN). The TGN is a fully-automated listening and speaking test designed to measure the ability of a test taker to understand basic Dutch spoken at a normal conversational speed, and to respond correctly and intelligibly in Dutch.

The TGN test, which takes about twelve minutes to complete, is administered via the telephone and scored within minutes by the patented Ordinate® automated spoken language assessment technology. This technology was designed to not only measure spoken content and structure, but also to evaluate pronunciation and fluency.

The TGN test is only administered in certified locations such as consulates and embassies of The Netherlands using landline telephones. Prior to the test, test-takers receive verbal instructions in their native language or in another language that they understand.

During the test, the system presents 50 randomly recorded spoken prompts from a variety of Dutch speakers at a conversational pace, each of which elicits an oral response in Dutch. The questions do not presume

any knowledge of specific facts of Dutch or Germanic culture, geography, history, or other subject matter. Each question is within the realm of familiarity of both a typical 12-year-old native speaker of Dutch and/or an adult who has never lived in a Dutch-speaking country.

The test is organized into five sections:

- Section A: Repeats (Part 1)
- Section B: Short-Answer Questions
- Section C: Repeats (Part 2)
- Section D: Opposites
- Section E: Story Retellings

Test takers are introduced to each type of question and then simply listen to the prompts and respond. Once the test is over, the results can be accessed online within a few minutes, using the unique test identification number (TIN) assigned to each test.

Results: Successfully Integrated TGN Test into Dutch Embassies and Consulates

As soon as the test was ready, and before it was deployed, the Ministry of Justice conducted an evaluation to ensure its consistency and accuracy. The Ministry reviewed all elements of the test including administration conditions, content of the questions/items, test language, scoring, scaling and norming, test reliability, and methods for fraud prevention.

In its final report, the Ministry concluded that TGN is a valid test of spoken Dutch. They noted that “native speakers of Dutch...get maximum scores, while learners of Dutch as a second language get a broad distribution of scores.” In sum, the Ministry found that the relationship between actual speaking skill and the speaking skill demonstrated during the test is consistent, free of bias for country of origin, age, gender, and degree of literacy.



Testing Fairness

Fairness is critical for the success of any testing system. Pearson maintains an extremely thorough development and quality control process for every test.

- The TGN scores are benchmarked to the levels of the Common European Framework of Reference (CEF-R, Common European Framework of Reference for Languages, Council for Europe, 2001).
- Ratings from 145 human judges of the test subjects were used for scaling and norming.
- The questions were created with a view toward candidates with low levels of education. Questions which require basic arithmetic skills were also avoided.
- The scoring system was created using the speech of 1,500 learners of Dutch (from 121 countries) and of 800 Dutch native speakers.

For more details on the test validation, write to Pearson at sales@pearsonkt.com and request a copy of its *Dutch Oral Interaction Test: Test Description & Validation Summary*.

“Only Pearson could have created a test that is extremely accurate and easy to deploy. We get results fast with a test that is as accurate as human evaluators, but we save the immense time and expense required with deploying human evaluators around the world.”

– Petra Poelmans, PhD, Consultant with CINOP

Additionally, “The test reliability of the TGN in terms of homogeneity (the extent to which all items contribute to the same measurement) amounts to 0.94.” In other words, a test taker is likely to earn the same or similar score consistently even if they took the test on numerous occasions. Further, the machine-generated scores correlate with human judgments at 0.95, demonstrating that the TGN machine-generated scores are virtually indistinguishable from the scoring that is done by careful human transcription and repeated independent human judgments. Based on these empiric findings, the Ministry was able to confirm that the TGN, using the Ordinate automated spoken language assessment technology, was an acceptable replacement for human judges.

The TGN was successfully deployed to all Dutch embassies and consulates in 2005 and is now the official test of spoken Dutch for all immigration applicants, allowing the Ministry to meet the requirements set forth by the Dutch parliament.

Ordinate Testing Technology

Ordinate, creator of the Versant tests, has been combined with Pearson’s Knowledge Technologies group. Ordinate has been producing automated spoken language tests since 1996 and is the leader in the field. The patented Ordinate® automated testing system was developed to apply advanced speech recognition techniques and data collection via the telephone to the evaluation of language skills. The TGN is the result of years of research in speech recognition, statistical modeling, linguistics, and testing theory. Ordinate scoring technology has been applied primarily to Pearson’s own language tests, such as the TGN and the Versant series, but also to customized tests. Applications of Ordinate technology include assessment of spoken English, children’s reading proficiency, adult literacy assessment, and the collection and human rating of spoken language samples.



Contact Us

To try a sample test or get more information, contact us at:

US: 800.211.8378
Int'l: +1 650.470.3505
sales@pearsonkt.com

Or visit us online at:
www.VersantTest.com