



Versant™ French Test

Test Description and Validation Summary

Table of Contents

1. Introduction	3
2. Test Description	3
2.1 Test Design	3
2.2 Test Administration	4
2.2.1 Telephone Administration	4
2.2.2 Computer Administration	4
2.3 Test Format	4
Part A: Reading	4
Part B: Repeats	5
Part C: Short Answer Questions	6
Part D: Sentence Builds	6
Part E: Story Retelling	6
Part F: Open Questions	7
2.4 Number of Items	7
2.5 Test Construct	8
3. Content Design and Development	10
3.1 Vocabulary Selection	11
3.2 Item Development	12
3.3 Item Prompt Recordings	12
3.3.1 Voice Talent for Test Items	12
3.3.2 Recording Review	12
4. Score Reporting	13
4.1 Scores and Weights	13
4.2 Score Use	15
4.3 Score Interpretation	15
5. Validation	15
5.1 Validity Study Design	16
5.1.1 Validation Sample	16
5.2 Internal Validity	16
5.2.1 Standard Error of Measurement	16
5.2.2 Reliability	17
5.2.3 Dimensionality: Correlation between Subscores	17
5.2.4 Correlations between the Versant French Test and Human Scores	18
5.3 Relationship to Known Populations: Native and Non-native Group Performance	19
5.4 Relationship to Scores of Tests with Related Constructs	20
6. Conclusions	21
7. About the Company	22
8. References	22
9. Appendix A: Test Paper	25
10. Appendix B: CEFR Mapping	27

1. Introduction

The Versant™ French Test, powered by Ordinate technology, is an assessment instrument designed to measure how well a person understands and speaks French. The Versant French Test is intended for adults and students over the age of 15 and takes approximately 16 minutes to complete. Because the Versant French Test is delivered automatically by the Versant testing system, the test can be taken at any time, from any location by phone or via computer. A human examiner is not required. The computerized scoring allows for immediate, objective, and reliable results that correspond well with traditional measures of spoken French performance.

The Versant French Test measures *facility* with spoken French, which is a key element in French oral proficiency. Facility in spoken French is how well the person can understand spoken French on everyday topics and respond appropriately at a native-like conversational pace in French. Academic institutions, corporations, and government agencies may use the Versant French Test to evaluate the ability of students, staff, and officers to understand spoken French and to express themselves clearly and appropriately in French. Scores from the Versant French Test provide reliable information that can be applied to placement, qualification and certification decisions, as well as monitor progress and measure instructional outcomes.

2. Test Description

2.1 Test Design

The Versant French Test may be taken at any time from any location using a telephone or a computer. During test administration, the Versant testing system presents a series of recorded spoken prompts in French at a conversational pace and elicits oral responses in French. The voices of the item prompts are from native speakers of French from Canada, France, and Africa, providing a range of speaking styles and accents.

The Versant French Test has six item types: Reading, Repeats, Short Answer Questions, Sentence Builds, Story Retelling, and Open Questions. All item types except for Open Questions elicit responses that can be analyzed automatically. These item types provide multiple, fully independent measures that underlie facility with spoken French, including phonological fluency, sentence construction and comprehension, passive and active vocabulary use, listening skill, and pronunciation of rhythmic and segmental units. Because more than one item type contributes to each subscore, the use of multiple item types strengthens score reliability.

The Versant testing system analyzes the candidate's responses and posts scores to a secure website usually within minutes of the completed test. Test administrators and score users can view and print out test results from a password-protected website.

The Versant French Test provides numeric scores and performance levels that describe the candidate's facility in spoken French – that is, the ability to understand spoken French on everyday topics and to respond appropriately at a native-like conversational pace in intelligible French. The Versant French Test score report is comprised of an Overall score and four diagnostic subscores: Sentence Mastery, Vocabulary, Fluency, and Pronunciation. Together, these scores describe the candidate's facility in spoken French.

2.2 Test Administration

Administration of a Versant French Test generally takes about 16 minutes over the telephone or via a computer. It is recommended (even for computer delivered tests) for the administrator to give a test paper to the candidate at least five minutes before starting the test (see Appendix). The candidate then has the opportunity to read both sides of the test paper and ask questions before the test begins. The administrator should answer any procedural or content questions that the candidate may have.

The delivery of the recorded item prompts is interactive – the system detects when the candidate has finished responding to one item and then presents the next item.

2.2.1 Telephone Administration

Telephone administration is supported by a test paper. The test paper is a single sheet of paper with material printed on both sides. The first side contains general instructions and an explanation of the test procedures. These instructions are the same for all candidates. The second side has the individual test form, which contains the phone number to call, the Test Identification Number (TIN), the spoken instructions written out verbatim, item examples, and the printed sentences for Part A: Reading. The individual test form is unique for each candidate.

When the candidate calls the Versant testing system, the system will ask the candidate to use the telephone keypad to enter the Test Identification Number that is printed on the test paper. This identification number is unique for each candidate and keeps the candidate's information secure.

A single examiner voice presents all the spoken instructions for the test. The spoken instructions for each section are also printed verbatim on the test paper to help ensure that candidates understand the directions. These test instructions are available in French, English, and Spanish. Candidates interact with the test material in French, going through all six parts of the test until they complete the test and hang up the telephone.

2.2.2 Computer Administration

For computer administration, the computer must have an Internet connection and Pearson's Computer Delivered Test (CDT) software (available at <http://www.versanttest.com/technology/platforms/cdt/index.jsp>). The candidate is fitted with a microphone headset. The CDT software prompts the candidate to adjust the volume and calibrate the microphone before the test begins.

The instructions for each section are spoken by an examiner voice and are also displayed on the computer screen. Candidates interact with the test system in French, speaking their responses into the microphone. When a test is finished, the candidate clicks a button labeled, "End Test".

2.3 Test Format

The following subsections provide brief descriptions of the item types and the abilities required to respond to the items in each of the six parts of the Versant French Test.

Part A: Reading

In this task, the candidate reads printed, numbered sentences, one at a time, as prompted. For telephone administration, the sentences are printed on the test paper. For computer administration, the sentences are displayed on the computer screen. Reading items are grouped into sets of four sequentially coherent sentences, as in the examples below.

Example:

1. Margot et sa meilleure amie se querellent souvent pour des riens.
2. La dernière fois, Margot s'est moquée de l'acteur favori de son amie.
3. Son amie s'est fâchée très vite.
4. Mais elle a vu que Margot ne faisait que la taquiner et elle a bien ri!

Presenting the sentences as part of a group helps the candidate disambiguate words in context and helps suggest how each individual sentence should be read aloud. The computer screen or test paper contains three groups of four sentences (i.e., 12 items). Candidates are prompted to read eight of the twelve sentences in a random order. The system tells the candidate which of the numbered sentences to read aloud (e.g., "Now, please read sentence 7."). After the candidate has read the sentence (or has remained silent for a period of time), the system prompts him or her to read another sentence from the list.

The sentences are relatively simple in structure and vocabulary, so literate speakers of French can read them easily and in a fluent manner. For candidates with little facility in spoken French but with some reading skills, this task provides samples of their pronunciation and reading fluency. The readings appear first in the test because, for many candidates, reading aloud presents a familiar task and is a comfortable introduction to the interactive mode of the test as a whole.

Part B: Repeats

In this task, candidates are asked to repeat sentences that they hear verbatim. The sentences are presented to the candidate in approximate order of increasing difficulty. Sentences range in length from three words to 15 words. The audio item prompts are spoken in a conversational manner.

Examples:

- La place est vide.
La chaleur est étouffante aujourd'hui !
Il faut se dépêcher, le train part dans dix minutes.

To repeat a sentence longer than about seven syllables, a person must recognize the words as spoken in a continuous stream of speech (Miller & Isard, 1963). Highly proficient speakers of French can generally repeat sentences that contain many more than seven syllables because these speakers are very familiar with French words, phrase structures, and other common syntactic forms. If a person habitually processes five-word or six-word phrases as one unit (e.g., "la plus grande salle de cinéma"), then that person can usually repeat utterances of 15 or 20 words in length without difficulty. Generally, the ability to repeat material is constrained by the size of the linguistic unit that a person can process in an automatic or nearly automatic fashion. As the sentences increase in length and complexity, the task becomes increasingly difficult for speakers who are not familiar with French sentence structure.

Because the Repeat items require candidates to organize speech into linguistic units, Repeat items assess the candidate's mastery of phrase and sentence structure. Given that the task requires the candidate to repeat full sentences (as opposed to just words and phrases), it also offers a sample of the candidate's fluency and pronunciation in continuous spoken French.

Part C: Short Answer Questions

In this task, candidates listen to spoken questions and answer each question with a single word or short phrase. The questions generally present at least three or four lexical items spoken in a continuous phonological form and framed in French sentence structure. Each question asks for basic information or requires simple inferences based on time, sequence, number, lexical content, or logic. The questions do not presume any knowledge of specific facts of culture, geography, history, or other subject matter; they are intended to be within the realm of familiarity of both a typical 12-year-old native speaker of French and an adult who has never lived in a French-speaking country.

Examples:

Combien de jours dans une semaine ?
Où porte-t-on des bottes : aux pieds ou dans le dos ?
Utilise-t-on un balai ou un peigne pour se démêler les cheveux ?

To correctly respond to the questions, a candidate must identify the words in phonological and syntactic context, and then infer the demand proposition. Short Answer Questions measure receptive and productive vocabulary within the context of spoken questions presented in a conversational style.

Part D: Sentence Builds

For the Sentence Builds task, candidates hear three short phrases and are asked to rearrange them to make a sentence. The phrases are presented in a random order (excluding the original word order), and the candidate says a reasonable and grammatical sentence that comprises exactly the three given phrases.

Examples:

vite / plus / cours
souvent d'accord / ne sont pas / les diplomates
des pluies abondantes / les météorologues avaient prédit / dans toutes les régions

To correctly complete this task, a candidate must understand the possible meanings of the phrases and know how they might combine with other phrasal material, both with regard to syntax and pragmatics. The length and complexity of the sentence that can be built is constrained by the size of the linguistic unit (e.g., a one word versus a three-word phrase) that a person can hold in verbal working memory. This is important to measure because it reflects the candidate's ability to access and retrieve lexical items and to build phrases and clause structures automatically. The more automatic these processes are, the more the candidate's facility in spoken French. This skill is demonstrably distinct from memory span (see Section 2.5, Test Construct, below).

The Sentence Builds task involves constructing and articulating entire sentences. As such, it is a measure of candidates' mastery of sentences in addition to their pronunciation and fluency.

Part E: Story Retelling

In this task, candidates listen to a brief story and are then asked to describe what happened in their own words. Candidates have 30 seconds to respond to each story. Candidates are encouraged to tell as

much of the story as they can, including the situation, characters, actions and ending. The stories consist of three to six sentences and contain from 30 to 90 words. The situation involves a character (or characters), setting, and goal. The body of the story describes an action by the agent of the story followed by a possible reaction or implicit sequence of events. The ending typically introduces a new situation, actor, patient, thought, or emotion.

Examples:

Deux petites filles jouaient au ballon sur la plage quand le vent s'est mis à souffler et a envoyé leur ballon dans la mer. Heureusement, un nageur a vu le ballon et le leur a rapporté. Pour le remercier, elles lui ont donné un beau coquillage.

Margot voulait aller étudier à l'étranger pour compléter ses études universitaires, mais elle ne savait pas quel pays choisir. Son amie lui a suggéré d'aller en Espagne, mais Margot ne parlait pas espagnol. Son professeur lui a recommandé d'aller en Angleterre pour améliorer son anglais, mais quand elle s'est finalement décidée de s'inscrire, il était trop tard : elle avait manqué la date limite d'inscription au programme. Déçue, elle a abandonné son projet et décidé de terminer ses études en France et de voyager après sa graduation.

The Story Retelling items assess a candidate's ability to listen and understand a passage, reformulate the passage using his or her own vocabulary and grammar, and then retell it in detail. This section elicits longer, more open-ended speech samples than earlier sections in the test, and allows for the assessment of a wider range of spoken abilities. Performance on Story Retelling provides a measure of fluency, pronunciation, vocabulary, and sentence mastery.

Part F: Open Questions

In this task, candidates listen to spoken questions that elicit an opinion, and are asked to provide an answer with an explanation. Candidates have 40 seconds to respond to each question. The questions relate to day-to-day issues or ask about the candidate's preferences and choices.

Example:

Croyez-vous qu'il soit important de faire de l'exercice régulièrement? Expliquez pourquoi.

D'après vous, est-il mieux d'apprendre une langue étrangère quand on est enfant ou adulte? Donnez vos raisons.

This task is used to collect longer spontaneous speech samples. Candidates' responses to items in this section are not scored, but are available for review by authorized listeners.

2.4 Number of Items

In the administration of the Versant French Test, the testing system serially presents a total of 63 items in six separate sections to each candidate. The 63 items are drawn at random from a large item pool. For example, each candidate is presented with 10 Sentence Builds from among those items available in

the pool, so most or all items will be different from one test administration to the next. Proprietary algorithms are used by the testing system to select from the item pool – the algorithms take into consideration, among other things, an item’s difficulty level and similarity to other presented items. Table I shows the number of items presented in each section.

Table I. Number of items presented per section.

Task	Presented
A. Reading	8
B. Repeat	16
C. Short Answer Questions	24
D. Sentence Builds	10
E. Story Retelling	3
F. Open Questions	2
Total	63

2.5 Test Construct

For any language test, it is essential to define the test construct as explicitly as possible (Bachman, 1990; Bachman & Palmer, 1996). The Versant French Test is designed to measure a candidate's facility in spoken French – that is, the ability to understand spoken French on everyday topics and to respond appropriately at a native-like conversational pace in intelligible French. Another way to describe the construct *facility in spoken French* is “the ease and immediacy in understanding and producing appropriate conversational [French]” (Levelt, 1989). This definition relates to what occurs during the course of a spoken conversation. While keeping up with the conversational pace, a person has to track what is being said, extract meaning as speech continues, and then formulate and produce a relevant and intelligible response. These component processes of listening and speaking are schematized in Figure 1.

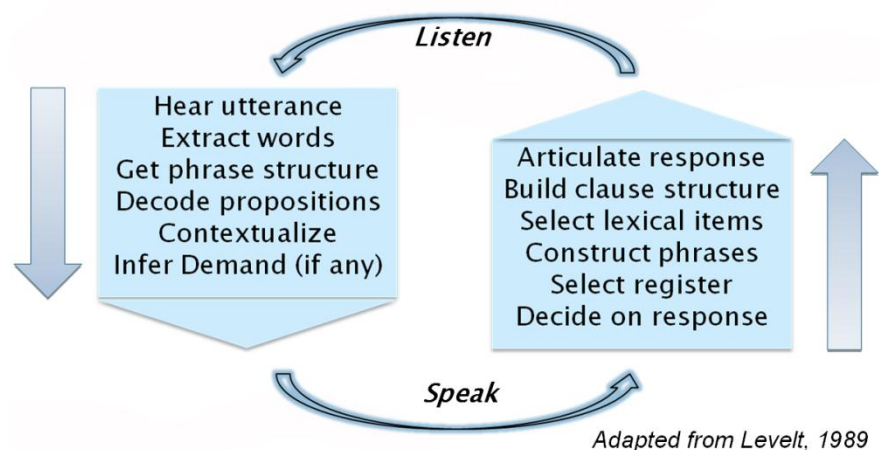


Figure 1. Conversational processing components in listening and speaking.

During a test, the testing system presents a series of discrete prompts to the candidate at a conversational pace as recorded by several different native speakers who represent a range of native accents and speaking styles. These integrated “listen-then-speak” items require real-time receptive and

productive processing of spoken language forms. The items are designed to be relatively independent of social nuance and higher cognitive functions. The same facility in spoken French that enables a person to participate in everyday native-paced French conversation also enables that person to satisfactorily understand and respond to the listening/speaking tasks in the Versant French Test.

The Versant French Test measures the candidate's control of core language processing components, such as lexical access and syntactic encoding. For example, in normal everyday conversation, native speakers go from building a clause structure to phonetic encoding (the last two stages in the right-hand column of Figure 1) in about 40 milliseconds (Van Turenhout, Hagoort, & Brown, 1998). Similarly, the other stages shown in Figure 1 must be performed within the short period of time available to a speaker during a conversational turn in everyday communication. The typical time window in turn taking is about 500-1000 milliseconds (Bull & Aylett, 1998). If language users involved in communication cannot successfully perform the complete series of mental activities presented in Figure 1 in real-time, both as listeners and as speakers, they will not be able to participate actively in conversations and other types of communication.

Automaticity in language processing is required in order for the speaker/listener to be able to pay attention to what needs to be said/understood rather than to how the encoded message is to be structured/analyzed. Automaticity in language processing is the ability to access and retrieve lexical items, to build phrases and clause structures, and to articulate responses without conscious attention to the linguistic code (Cutler, 2003; Jescheniak, Hahne, & Schriefers, 2003; Levelt, 2001). Some measures of automaticity in the Versant French Test may be misconstrued as memory tests. Because some tasks involve repeating long sentences or holding phrases in memory in order to piece them together into reasonable sentences, it may seem that these tasks are measuring memory capacity rather than language ability. However, psycholinguistic research has shown that verbal working memory for such things as remembering a string of digits is distinct from the cognitive resources used to process and comprehend sentences (Caplan & Waters, 1999).

The fact that syntactic processing resources are generally separate from short-term memory stores is also evident in the empirical results of the Versant French Test validation experiments (see Section 5: Validation). Virtually all native French speakers achieve high scores on the Versant French Test, whereas non-native speakers obtain scores distributed across the scale. If memory, as such, were being measured as an important component of performance on the Versant French Test, then native speakers would show greater variation in scores as a function of their range of memory capacities. The Versant French test would not correlate as highly as it does with other accepted measures of oral proficiency, since it would be measuring something other than language ability.

The Versant French Test probes the psycholinguistic elements of spoken language performance rather than the social, rhetorical, and cognitive elements of communication. The reason for this focus is to ensure that test performance relates most closely to the candidate's facility with the language itself and is not confounded with other factors. The goal is to separate familiarity with spoken language from other types of knowledge including cultural familiarity, understanding of social relations and behavior, and the candidate's own cognitive style. Also, by focusing on context-independent material, less time is spent developing a background cognitive schema for the tasks, and more time is spent collecting data for language assessment (Downey et al., 2008).

The Versant French Test measures the real-time encoding and decoding of spoken French. Performance on Versant French Test items predicts a more general spoken language facility, which is essential in successful oral communication. The reason for the predictive relation between spoken language facility and oral communication skills is schematized in Figure 2. This figure puts Figure 1 into a larger context, as one might find in a social-situated dialog.

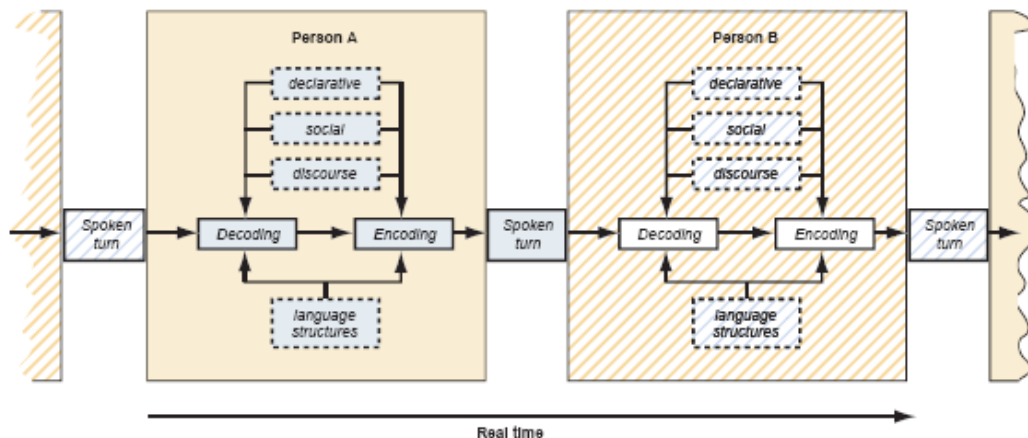


Figure 2. Message decoding and message encoding as a real-time chain-process in oral interaction.

The language structures that are largely shared among the members of a speech community are used to encode and decode various threads of meaning that are communicated in spoken turns. These threads of meaning that are encoded and decoded include declarative information, as well as social information and discourse markers. World knowledge and knowledge of social relations and behavior are also used in understanding and in formulating the content of the spoken turns. However, these social-cognitive elements of communication are not represented in this model and are not directly measured in the Versant French Test.

3. Content Design and Development

The Versant French Test measures both listening and speaking skills, emphasizing the candidate's facility (ease, fluency, immediacy) in responding aloud to common, everyday spoken French. All Versant French Test items are designed to be region neutral. The content specification also requires that both native speakers and highly proficient non-native speakers find the items very easy to understand and to respond to appropriately. For French learners, the items cover a broad range of skill levels and skill profiles.

Except for the Reading items, each Versant French Test item is independent of the other items and presents unpredictable spoken material in French. The test is designed to use context-independent material for three reasons. First, context-independent items exercise and measure the most basic meanings of words, phrases, and clauses on which context-dependent meanings are based (Perry, 2001). Second, when language usage is relatively context-independent, task performance depends less on factors such as world knowledge and cognitive style and more on the candidate's facility with the language itself. Thus, the test performance on the Versant French Test relates most closely to language abilities and is not confounded with other candidate characteristics. Third, context-independent tasks maximize response density; that is, within the time allotted, the candidate has more time to demonstrate performance in speaking the language. Less time is spent developing a background cognitive schema needed for successful task performance. Item types maximize reliability by providing multiple, fully independent measures. They elicit responses that can be analyzed automatically to produce measures that underlie facility with spoken French, including phonological fluency, sentence comprehension, vocabulary, and pronunciation of lexical and phrasal units.

3.1 Vocabulary Selection

The vocabulary used in all test items and responses is restricted to forms of the 8,000 most frequently occurring words found in the Versant French Test Corpus- a large corpus compiled by Pearson from six unique sources. The sources were chosen to represent spontaneous speech and conversations, as well as common written forms. They were also chosen to capture the variety of lexical items found in various regions of France and Canada, and represent speakers of both sexes from diverse ages and socio-economic backgrounds. Table 2 below lists and describes the nature of each of the six sources. The Versant French Test Corpus was compiled as follows: for each source, the most frequent 10,000 words was captured via computer frequency search and entered into a database. (The process of first calculating frequency by source assured that each source was given equal weight in the selection process and that sources with a larger number of words did not dominate the pool.) Once these 48,000 words were compiled from each source, the most frequent 8,000 words among them were identified to form the final list of lexical forms.

Table 2. Sources used to create the Versant French Test Corpus

Source	Source Description	Source type	Timeframe	Region	Total # of Words
Journal des débats de l'Assemblée nationale (Assemblée nationale du Québec, 2010)	Proceedings of the Canadian National Assembly	Spontaneous and rehearsed political speeches and debates	1989-2009	Canada	4.3 million
Dictionnaire De Frequence Des Mots Du Francaise Parle Au Quebec (Beauchemin, Martel, & Theoret, 1992)	Frequency dictionary compiled from transcribed conversations	Spontaneous casual conversations between friends	1970-1979	Throughout Canada	11,000 most frequent compiled from over 1 million
European Parliament Proceedings Parallel Corpus (Koehn, 2005)	Proceedings of European Parliament	Spontaneous and rehearsed political speeches and debates	1996-2009	France	29 million
Corpus d'Orléans (Baude, 2010)	Conversations between people of various ages and socio-economic backgrounds	Spontaneous casual conversations between friends	Late 1968-1971	Orleans, France	217,000
Bristol Corpus (Beeching, 2001)	Conversations and interviews with a wide ranges of age groups	Spontaneous casual conversation and informal interviews	1980-1990	Northern and Southern France	155,000
A Frequency Dictionary of French: Core Vocabulary for Learners	5000 most frequent words compiled from spoken and written sources	Conversations, interviews, parliamentary debates, film subtitles, books,	1950-2005	"Both France and overseas"	Most frequent 5000 compiled from 11.5 million spoken

(Lonsdale & Le Bras, 2009)		magazines, and newspapers.			and 11.5 million written
----------------------------	--	----------------------------	--	--	--------------------------

3.2 Item Development

Versant French Test items were drafted by native French-speaking item developers from different regions of France and Canada. In general, the language structures used in the test reflect those that are common in everyday conversational French familiar to native speakers from both Canada and France. The items were designed to be independent of social nuance and complex cognitive functions.

Draft items underwent several rounds of review, both internal and external. First, all items written by Canadian item writers were reviewed by a European French item writer and all items written by European French item writers were reviewed by a Canadian item writer. Items were then sent back to the original author to be revised. The process was repeated until the item was deemed acceptable by both writer and reviewer. Items were then reviewed internally by a team of test developers, all with advanced degrees in language-related fields, to ensure that they conformed to item specifications and contained appropriate content. Then, draft items were sent to external linguists (university professors in Canada and France) for expert review to ensure 1) compliance with the vocabulary specification, and 2) conformity with current colloquial French usage in different countries. Reviewers checked that items would be appropriate for candidates trained to standards in either European or Canadian French.

All items, including anticipated responses for short-answer questions, were also checked for compliance with the vocabulary specification. Most vocabulary items that were not present in the lexicon were changed to other lexical stems that were in the consolidated word list. Some off-list words were kept and added to a supplementary vocabulary list, as deemed necessary and appropriate. Changes proposed by the different reviewers were then reconciled and the original items were edited accordingly.

For an item to be retained in the test, it had to be understood and responded to appropriately by at least 85% of a reference sample of educated native speakers of Canadian or European French.

3.3 Item Prompt Recordings

3.3.1 Voice Talent for Test Items

Fifteen native speakers (6 men and 8 women) representing various speaking styles, ages, and regions of France (8 speakers), Canada (4 speakers), and Africa (3 speakers) were selected for recording the spoken prompt materials (all item types except Reading items). These speakers were asked to record the items in a clear, natural, and conversational tone.

Recordings were made in a professional recording studio in Menlo Park, California. In addition to the item prompt recordings, all the test instructions were recorded by a professional male voice talent whose voice is distinct from the item voices.

3.3.2 Recording Review

Multiple independent reviews by French linguists and language assessment experts were performed on all the recordings for quality, clarity, and conformity to natural conversational styles. Any recording in which reviewers noted some type of error was either re-recorded or excluded from insertion in the operational test.

4. Score Reporting

4.1 Scores and Weights

The Versant French Test score report is comprised of an Overall score and four diagnostic subscores (Sentence Mastery, Vocabulary, Fluency¹ and Pronunciation).

Overall: The Overall score of the test represents the ability to understand spoken French and speak French intelligibly at a native-like conversational pace on everyday topics. Scores are based on a weighted combination of the four diagnostic subscores. Scores are reported in the range from 20 to 80.

Sentence Mastery: Sentence Mastery reflects the ability to understand, recall, and produce French phrases and clauses in complete sentences. Performance depends on accurate syntactic processing and appropriate usage of words, phrases, and clauses in meaningful sentence structures.

Vocabulary: Vocabulary reflects the ability to understand common everyday words spoken in sentence context and to produce such words as needed. Performance depends on familiarity with the form and meaning of everyday words and their use in connected speech.

Fluency: Fluency is measured from the rhythm, phrasing and timing evident in constructing, reading and repeating sentences.

Pronunciation: Pronunciation reflects the ability to produce consonants, vowels, and stress in a native-like manner in sentence context. Performance depends on knowledge of the phonological structure of everyday words as they occur in phrasal context.

Of the 63 items in an administration of the Versant French Test, 57 responses are currently used in the automatic scoring. The first item response in Parts A through D is considered a practice item and is not incorporated into the final score. The two Open Questions are not scored. Figure 3 illustrates which sections of the test contribute to each of the four subscores. Each vertical rectangle represents a response from a candidate. The items that are not included in the automatic scoring are shown in blue.

¹ Within the context of language acquisition, the term “fluency” is sometimes used in the broader sense of general language mastery. In the narrower sense used in the Versant French Test score reporting, “fluency” is taken as a component of oral proficiency that describes certain characteristics of the observable performance. Following this usage, Lennon (1990) identified fluency as “an impression on the listener’s part that the psycholinguistic processes of speech planning and speech production are functioning easily and efficiently” (p. 391). In Lennon’s view, surface fluency is an indication of a fluent process of encoding. The Versant French Test fluency subscore is based on measurements of surface features such as the response latency, speaking rate, and continuity in speech flow, but as a constituent of the Overall score it is also an indication of the ease of the underlying encoding process.

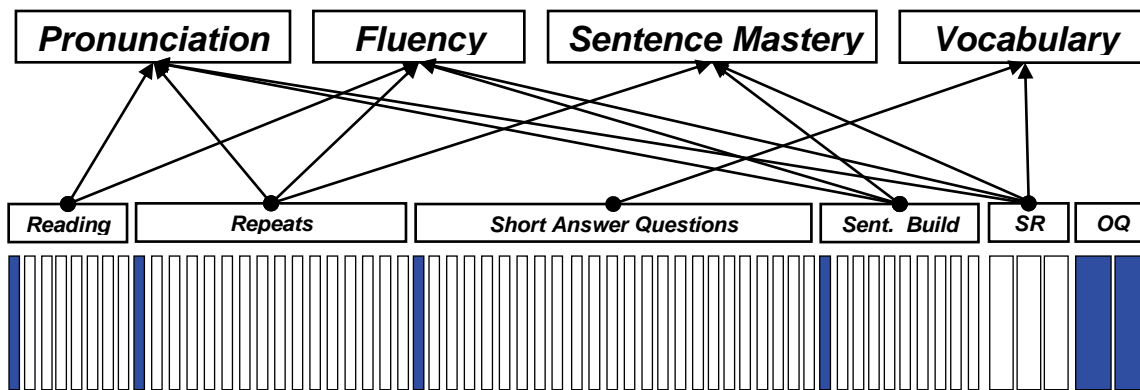


Figure 3. Relation of subscores to item types.

Among the four subscores, two basic types of scores are distinguished: scores relating to the content of what a candidate says (Sentence Mastery and Vocabulary) and scores relating to the manner (quality) of the response production (Fluency and Pronunciation). These two types of scores correspond roughly to Carroll's (1961) distinction between a knowledge aspect and a control aspect of language performance. In later publications, Carroll (1986) identified the control aspect as automatization, which suggests that people speaking fluently without realizing they are using their knowledge about a language have attained a level of automatic processing.

In all but the Open Questions section of the Versant French Test, each incoming response is recognized automatically by a speech recognizer that has been optimized for non-native speech. The words, pauses, syllables, phones, and even some subphonemic events are located in the recorded signal. The content of the responses to Reading, Repeats, SAQs, and Sentence Builds is scored according to the presence or absence of expected correct words in correct sequences. The content of responses to Story Retelling items is scored for vocabulary by scaling the weighted sum of the occurrence of a large set of expected words and word sequences that are recognized in the spoken response. Weights are assigned to the expected words and word sequences according to their semantic relation to the story prompt using a variation of latent semantic analysis (Landauer et al., 1998). Across all the items, content accuracy counts for 50% of the Overall score, and reflects whether or not the candidate understood the prompts and responded with appropriate content.

The manner-of-speaking scores (Fluency and Pronunciation, or the control dimension) are calculated by measuring the latency of the response, the rate of speaking, the position and length of pauses, the stress and segmental forms of the words, and the pronunciation of the segments in the words within their lexical and phrasal context. These measures are scaled according to the native and non-native distributions and then re-scaled and combined so that they optimally predict human judgments on manner-of-speaking. The manner-of-speaking scores count for the remaining 50% of the Overall score, and reflect whether or not the candidate speaks in a native-like manner.

In the Versant French Test scoring logic, content and manner (i.e. accuracy and control) are weighted equally because successful communication depends on both. Producing accurate lexical and structural content is important, but excessive attention to accuracy can lead to disfluent speech production and can also hinder oral communication; on the other hand, inappropriate word usage and misunderstood syntactic structures can also hinder communication.

4.2 Score Use

Once a candidate has completed a test, the Versant testing system analyzes the spoken performances and posts the scores at www.VersantTest.com. Test administrators and score users can then view and print out the test results from a password-protected section of the website.

Scores from the Versant French Test are intended for use by educational and government institutions as well as commercial and business organizations. Pearson endorses the use of Versant French Test scores for making valid decisions about oral French interaction skills of individuals, provided score users have reliable evidence confirming the identity of the individuals at the time of test administration. Score users may obtain such evidence either by administering *the* Versant French Test themselves or by having trusted third parties administer the test. In several countries, education and commercial institutions provide such services.

Versant French Test scores can be used to evaluate the level of spoken French skills of individuals entering into, progressing through, and exiting French language courses. Scores may also be used effectively in evaluating whether an individual's level of spoken French is sufficient to perform certain tasks or functions requiring mastery of spoken French.

The Versant French Test score scale covers a wide range of abilities in spoken French communication. In most cases, score users must decide what Versant French Test score is considered a minimum requirement in their context (i.e., a cut score). Score users may wish to base their selection of an appropriate cut score on their own localized research. Pearson can provide a Benchmarking Kit and further assistance in establishing cut scores.

4.3 Score Interpretation

Two summary tables offer a quick reference for interpreting Versant French Test scores in terms of the Common European Framework of Reference descriptors. Appendix B presents an overview relating the Common European Framework global scale (Council of Europe, 2001:24) to Versant French Test Overall scores. Table 7 in the Appendix provides the more specific scale of Oral Interaction Descriptors used in the studies designed to align the two scales. The method used to create the reference tables is described in a white paper. Please contact Pearson for this report.

5. Validation

The scoring models used in the Versant French Test were trained on a norming data set comprised of 291 native and 988 non-native French speaking test-takers. In this norming data set, the ages ranged from 16 to 74, with a mean age of 33. The female:male ratio was 64:36.

Within the norming data set, 440 took full-length tests; the data from these tests was used to train the scoring models for every item type. Of these 440 full-length test-takers, 100 were native French speakers: 52 from Canada, 47 from France, and 1 from Switzerland. The remaining 340 non-native test-takers came from 53 different countries from throughout North and South America, Eastern and Western Europe, The Middle East, Africa, and Asia. Their native languages included Albanian, Arabic, Bambara, Basaa, Bengali, Berber, Bicol, Bulaare, Cambodian, Cantonese, Catalan, Chinese, Creole, Dioula, Djerma, Douala, English, Farsi, French (reported by speakers who indicated French was their first language but who also spoke an African language from birth), Fulani, Fulfulde, German, Hausa, Hindi, Inuktitut, Italian, Japanese, Korean, Lingala, Mandarin, Mongolian, Persian, Portuguese, Romanian,

Russian, Spanish, Tagalog, Tamil, Thai, Vietnamese, Yoruba, and Zarma.

The other 839 test-takers in the norming data set were given tests that predominantly contained story retelling items; as this item-type requires substantial amounts of response data for automated model-building. Of this group, 191 were native French speakers: 108 from Canada, 77 from France, and 6 from Haiti. The remaining 648 non-native test-takers came from 57 countries from North and South America, Eastern and Western Europe, The Middle East, Africa, and Asia. Their native languages included Albanian, Arabic, Bambara, Basaa, Berber, Cantonese, Chinese, Creole, Djerma, Ekie, English, Farsi, French (reported by speakers who indicated French was their first language but who also spoke an African language from birth), Fulani, Fulfulde, German, Hausa, Hindi, Indonesian, Italian, Korean, Mandarin, Moldovan, Mossi, Nepalese, Persian, Polish, Portuguese, Romanian, Russian, Spanish, Tagalog, Tamil, Thai, Vietnamese, Xwla, Yoruba, and Zarma.

5.1 Validity Study Design

Validity analyses examined three aspects of the Versant French Test scores:

1. Internal quality (reliability and accuracy): whether or not the Versant French Test a) provides consistent scores that accurately reflect the scores that human listeners and raters would assign and b) provides distinct subscores that measure different aspects of the test construct.
2. Relation to known populations: whether or not the Versant French Test scores reflect expected differences and similarities among known populations (e.g., natives vs. French learners).
3. Relation to scores of tests with related constructs: how closely Versant French Test scores predict the reliable information in scores of well-established speaking tests.

5.1.1 Validation Sample

From the large body of spoken performance data collected from native and non-native speakers of French, a total of 150 participants were set aside for a series of validation analyses. Over 30 different languages from 26 countries were represented in the validation sample, including a total of five native French speakers. Ages ranged from 15 to 73 with a mean age of 33. The female:male ratio was 70:30. Care was taken to ensure that the training dataset and validation dataset did not overlap. That is, the spoken performance sample provided by the validation candidates were excluded from the datasets used for training the automatic speech processing models or for training any of the scoring models.

5.2 Internal Validity

To understand the consistency and accuracy of the Versant French Overall scores and the distinctness of the subscores, the following indicators were examined: the standard error of measurement of the Versant French Overall score; the reliability of the Versant French Test (split-half and test-retest); the correlations between the Versant French Overall scores and subscores, and between pairs of subscores; and comparison of machine-generated Versant French scores with listener-judged scores of the same Versant French tests. These qualities of consistency and accuracy of the test scores are the foundation of any valid test (Bachman & Palmer, 1996).

5.2.1 Standard Error of Measurement

The Standard Error of Measurement (SEM) provides an estimate of the amount of error in an individual's observed test scores and "shows how far it is worth taking the reported score at face value" (Luoma, 2004: 183). The SEM of the Versant French Overall score is 2.1.

5.2.2 Reliability

Split-half Reliability

Score reliabilities were estimated by the split-half method ($n=150$). Split-half reliability was calculated for the Overall score and all subscores. The split-half reliabilities use the Spearman-Brown Prophecy Formula to correct for underestimation and are similar to the reliabilities calculated for the uncorrected equivalent form dataset. The human scores were calculated from human transcriptions (for the Sentence Mastery and Vocabulary subscores) and human judgments (for the Pronunciation and Fluency subscores). Table 3 presents split-half reliabilities based on the same individual performances scored by careful human rating in one case and by independent automatic machine scoring in the other case. The values in Table 3 suggest that there is sufficient information in a Versant French Test item response set to extract reliable information, and that the effect on reliability of using the Ordinate speech recognition technology, as opposed to careful human rating, is quite small across all score categories. The high reliability score is a good indication that the computerized assessment will be consistent for the same candidate assuming no changes in the candidate's language proficiency level.

Table 3. Split-Half Reliabilities of Versant French Test Machine Scoring versus Human Scoring

Score	<i>Machine Split-half Reliability (n = 150)</i>	<i>Human Split-half Reliability (n=150)</i>
Overall	0.97	0.99
Sentence Mastery	0.89	0.93
Vocabulary	0.77	0.86
Fluency	0.93	0.99
Pronunciation	0.95	0.99

5.2.3 Dimensionality: Correlation between Subscores

Ideally, each subscore on a test provides unique information about a specific dimension of the candidate's ability. For spoken language tests, the expectation is that there will be a certain level of covariance between subscores given the nature of language learning. When language learning takes place, the candidate's skills tend to improve across multiple dimensions. However, if all the subscores were to correlate perfectly with one another, then the subscores might not be measuring different aspects of facility with the spoken language.

Table 4 presents the correlations among the Versant French Test subscores and the Overall scores for a semi-randomly selected non-native sample.

Table 4. Correlations among Versant French Test Subscores for a Semi-randomly Selected Non-Native Sample (n=330)

	Sentence Mastery	Vocabulary	Pronunciation	Fluency	Overall
Sentence Mastery	-	.78	.76	.70	.89
Vocabulary			.77	.72	.85
Pronunciation				.91	.95
Fluency					.93
Overall					-

As expected, test subscores correlate with each other to some extent by virtue of presumed general covariance within the candidate population between different component elements of spoken language skills. The correlations between the subscores are, however, significantly below unity, which indicates that the different scores measure different aspects of the test construct, using different measurement methods, and different sets of responses. This data set (n=330) was semi-randomly selected from the full length data collection tests. A broad range of native languages is represented. A different pattern may be found when different native languages are sampled.

5.2.4 Correlations between the Versant French Test and Human Scores

The final analysis for internal quality involved comparing scores from the Versant French Test using Pearson’s speech processing technologies versus careful human transcriptions and human judgments from expert raters. Table 5 presents correlations between machine-generated scores and human scores for the same subset of 150 candidates as given in section 5.2.2. The correlations presented in Table 5 suggest that the Versant French Test machine-generated scores are not only reliable, but that they generally correspond as they should with human ratings. Among the subscores, the human-machine relation is closer for the content accuracy scores than for the manner-of-speaking scores, but the relation is close for all four subscores.

Table 5. Correlations between the Versant French Test and Human Scores (n=150)

Score Type	Correlation
Overall	0.96
Sentence Mastery	0.97
Vocabulary	0.93
Fluency	0.88
Pronunciation	0.85

A scatterplot of human and machine scores for this subset is shown in Figure 4.

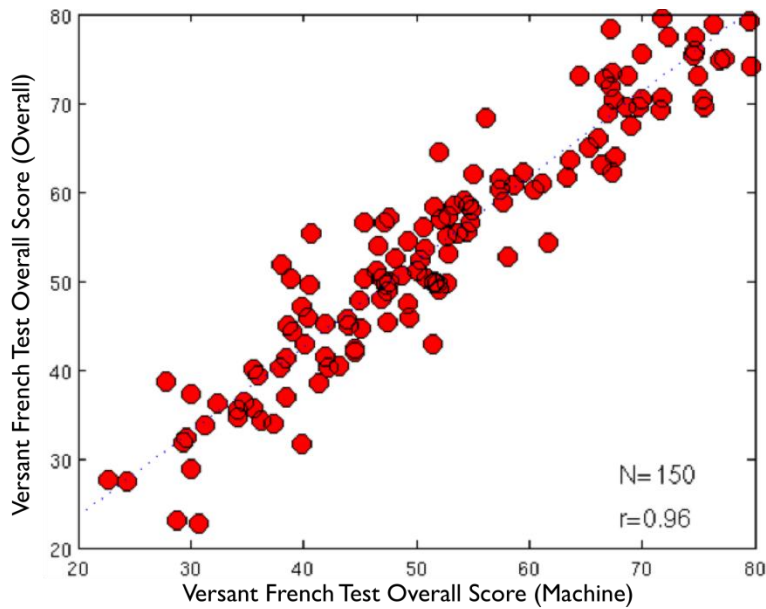


Figure 4. Versant French Test scores versus human scores ($n=150$).

In the scatterplot, all the data points fall within a tight range of the regression line with no outliers. Together the correlations and the scatterplot show that at the Overall score level, Versant French Test machine-generated scores are closely related scoring based on careful human transcriptions and repeated independent human judgments.

5.3 Relationship to Known Populations: Native and Non-native Group Performance

The next validity analysis examined whether or not the Versant French Test scores reflect expected differences between native and non-native French speakers. Overall scores from 88 native speakers and 330 non-native speakers representing a range of native languages were compared. Figure 5 presents cumulative distributions of Overall scores for the native and non-native speakers. Note that the range of scores displayed in this figure is from 10 through 90, whereas the Versant French Test scores are reported on a scale from 20 to 80. Scores outside the 20 to 80 range are deemed to have saturated the intended measurement range of the test and are therefore reported as 20 or 80.

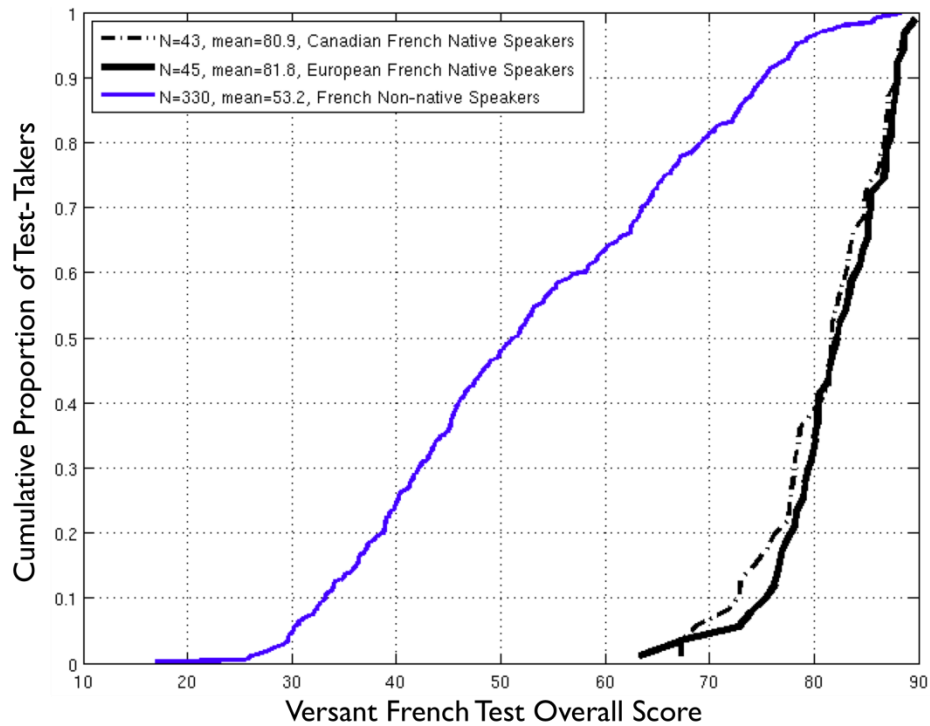


Figure 5. Cumulative density functions of Versant French Test Overall scores for the native and non-native norming groups (native $n=88$ and non-native $n=330$).

The results show that native speakers of French consistently obtain high scores on the Versant French Test. Fewer than 5% of the native sample scored below 70. Learners of French as a second or foreign language, on the other hand, are distributed over a wide range of scores. Note also that only 20% of the non-natives scored above 70. The Overall scores show effective separation between native and non-native candidates.

5.4 Relationship to Scores of Tests with Related Constructs

The Common European Framework of Reference (CEFR) is published by the Council of Europe, and provides a common basis for describing language proficiency using a six-level scale. In a study, six expert panelists independently evaluated 889 Story Retelling and Open Question responses from 185 unique test-takers using the CEFR descriptors. The correlation between a test-taker's Versant French Test overall score and his/her averaged assigned CEFR level was 0.88. Figure 6 illustrates the relation between the Overall scores on the Versant French Test and scores assigned by panelists using the CEFR. The graph shows how both instruments (Versant French Test and the CEFR) separate the native and non-native norming groups. The raters showed perfect agreement in assigning a Common European Framework (CEFR) level to 47% of the cases and differed by only one level in a further 45% of the cases. Rater agreement overall was 0.98. Final mappings between the two scales can be found in Appendix B.

Versant French Test Overall Score vs. CEFR Estimate

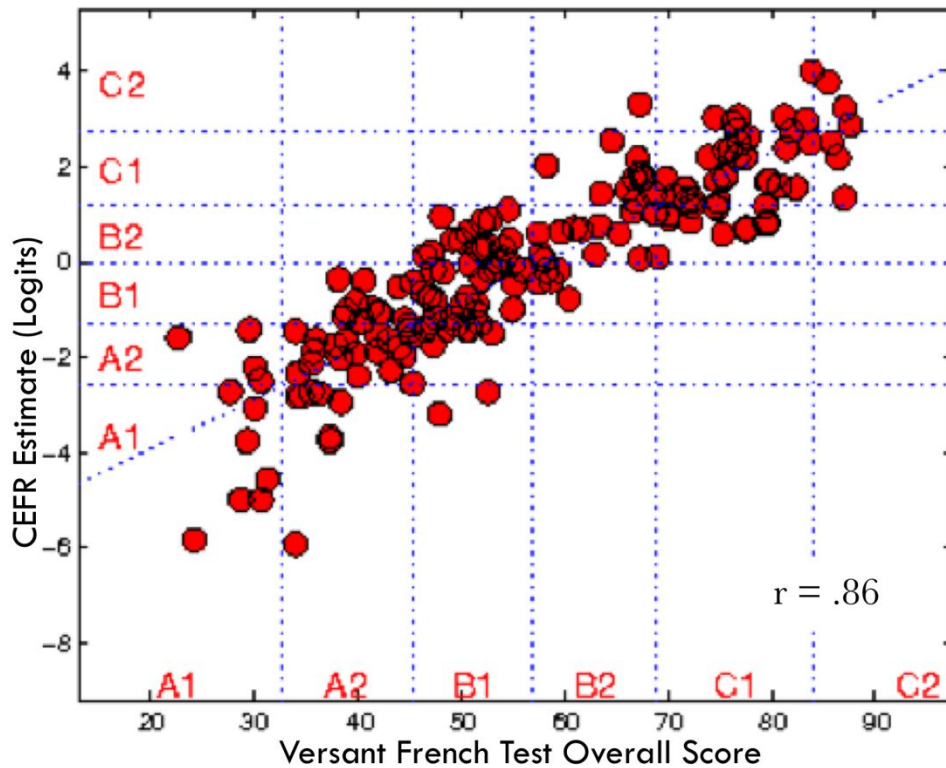


Figure 6. Correlation between Versant French Test Overall score and CEFR-levels (n=185).

6. Conclusions

Data from the validation studies provide evidence in support of the following conclusions:

- The Versant French Test produces precise and reliable skill estimates.
- Overall scores show effective separation between native and non-native candidates.
- Subscores of the Versant French Test are reasonably distinct and therefore offer useful diagnostics.
- Versant French Test scores show a high correlation with human-produced ratings.
- Versant French Test Overall scores have meaningful correlations with the CEFR scale as used for measuring French proficiency.

To assure the defensibility of employee selection procedures, employers in the U.S. follow the Equal Employment Opportunity Commission's (EEOC's) Uniform Guidelines for Employee Selection Procedures. These guidelines state that employee selection procedures must be reliable and valid. The above information provides evidence of the reliability, validity and legal defensibility of the Versant French Test in conformance with the prescriptions of the EEOC's Uniform Guidelines.

7. About the Company

Ordinate Testing Technology: The Versant automated testing system was developed to apply advanced speech recognition techniques and data collection to the evaluation of language skills. The system includes automatic telephone and computer reply procedures, dedicated speech recognizers, speech analyzers, databanks for digital storage of speech samples, and score report generators linked to the Internet. The *Versant French Test* is the result of years of research in speech recognition, statistical modeling, linguistics, and testing theory. The Versant patented technologies are applied to its own language tests such as the Versant series and also to customized tests. Sample projects include children's reading assessment, adult literacy assessment, and collections and human rating of spoken language samples.

Pearson: Pearson's Knowledge Technologies group and Ordinate Corporation, the creator of the Versant tests, were combined in January, 2008. The Versant tests are the first to leverage a completely automated method for assessing spoken language.

Pearson's Policy: Pearson is committed to the best practices in the development, use, and administration of language tests. Each Pearson employee strives to achieve the highest standards in test publishing and test practice. As applicable, Pearson follows the guidelines propounded in the Standards for Educational and Psychological Testing, and the Code of Professional Responsibilities in Educational Measurement. A copy of the Standards for Educational and Psychological Testing is available to every employee for reference.

Research at Pearson: In close cooperation with international experts, Pearson conducts ongoing research aimed at gathering substantial evidence for the validity, reliability, and practicality of its current products and investigating new applications for Ordinate technology. Research results are published in international journals and made available through the Versant website (www.VersantTest.com).

8. References

- Assemblée Nationale du Québec. (2010). Journal des débats de l'Assemblée nationale. Downloaded 9/2/10 from:
<http://www.assnat.qc.ca/fr/travaux-parlementaires/assemblee-nationale/39-1/journal-debats/20100325/13577.html>
- Bachman, L.F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L.F. & Palmer, A.S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Baude, O., compiler (2010.) Corpus d'Orléans. University Orléans. Downloaded 9/2/10 from National Library of Australia: <http://trove.nla.gov.au/book/result?l-language=fra&q=eslo&s=0>
- Beauchemin, N., Martel, P., & Theoret, M. (1992). *Dictionnaire de fréquence des mots du française parlé au Québec*. New York: Peter Lang Publishing.
- Beeching, K. (2001). *Le discours des hommes et des femmes dans le français contemporain. La fonction des incises parenthétiques et des particules énonciatives c'est-à-dire, enfin, hein et quoi*. Unpublished doctoral dissertation. University of Surrey.

- Bull, M & Aylett, M. (1998). An analysis of the timing of turn-taking in a corpus of goal-oriented dialogue. In R.H. Mannell & J. Robert-Ribes (Eds.), *Proceedings of the 5th International Conference on Spoken Language Processing*. Canberra, Australia: Australian Speech Science and Technology Association.
- Caplan, D. & Waters, G. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, 22, 77-126.
- Carroll, J.B. (1961). *Fundamental considerations in testing for French language proficiency of foreign students*. Testing. Washington, DC: Center for Applied Linguistics.
- Carroll, J.B. (1986). *Second language*. In R.F. Dillon & R.J. Sternberg (Eds.), *Cognition and Instructions*. Orlando FL: Academic Press.
- Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Council of Europe (2001). *Common European Framework of Reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Cutler, A. (2003). Lexical access. In L. Nadel (Ed.), *Encyclopedia of Cognitive Science*. Vol. 2, *Epilepsy – Mental imagery, philosophical issues about*. London: Nature Publishing Group, 858-864.
- Downey, R., Farhady, H., Present-Thomas, R., Suzuki, M., & Van Moere, A. (2008). Evaluation of the usefulness of the Versant for English test: A response. *Language Assessment Quarterly*, 5, 160-167.
- Equal Employment Opportunity Commission. *Uniform Guidelines for Employee Selection Procedures*. Retrieved from <http://www.uniformguidelines.com/uniformguidelines.html>
- Jescheniak, J.D., Hahne, A. & Schriefers, H.J. (2003). Information flow in the mental lexicon during speech planning: Evidence from event-related brain potentials. *Cognitive Brain Research*, 15(3), 261-276.
- Koehn, P. (2005). *Europarl: A parallel corpus for statistical machine translation*. MT Summit. Downloaded 9/1/10 from: <http://corp.hum.sdu.dk/index.html>
- Landauer, T.K., Foltz, P.W. & Laham, D. (1998). Introduction to Latent Semantic Analysis. *Discourse Processes*, 25, 259-284.
- Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, 40, 387-412.
- Levelt, W.J.M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Levelt, W.J.M. (2001). Spoken word production: A theory of lexical access. *PNAS*, 98(23), 13464-13471.

Lonsdale, D. & Le Bras, Y. (2009). *A frequency dictionary of French: Core vocabulary for young learners*. New York: Routledge.

Luoma, S. (2004). *Assessing Speaking*. Cambridge: Cambridge University Press.


Miller, G.A. & Isard, S. (1963). Some perceptual consequences of linguistic rules. *Journal of Verbal Learning and Verbal Behavior*, 2, 217-228.

Perry, J. (2001). *Reference and reflexivity*. Stanford, CA: CSLI Publications.

Van Turenhout, M., Hagoort, P. & Brown, C. M. (1998). Brain Activity During Speaking: From Syntax to Phonology in 40 Milliseconds. *Science*, 280, 572-574.

9. Appendix A: Test Paper




Side I of the Test Paper: Instructions and general introduction to test procedures. **Note:** These instructions are available in several different languages.



Test Instructions

Please read this before taking the test

Versant tests are automated spoken language tests that are taken on the telephone or computer. If you would like to listen to a sample test, purchase a practice test, or view the test score after taking the test (if applicable), please visit www.VersantTest.com

Part	Instructions
Before the Test	<ul style="list-style-type: none"> Carefully read this instruction page and the test paper. You may use a dictionary or ask someone for help if there are words or sentences that you don't understand. Choose a quiet location with a landline phone where you will not be interrupted during the test. Do not use a cordless phone, cellular phone, or VoIP phone (e.g., Skype™ or PC-to-phone services). Newer phones are generally better than older phones. Make sure that the phone is set to tone and not pulse.
Beginning the Test	<ul style="list-style-type: none"> To begin the test, call the phone number on the test paper using a landline push-button telephone. A recorded examiner's voice will guide you through each section of the test. Enter your Test Identification Number using the telephone keypad when the examiner's voice asks you to do so. This number is printed on the top right of your test paper. The examiner's voice will then ask you two questions: your name, and the city and the country you are calling from. If you are speaking too loudly or too quietly, the examiner's voice will tell you. The test begins when you say your name. If you hang up before you complete the test, the test cannot be graded. You cannot reuse the Test Identification Number.
During the Test	<ul style="list-style-type: none"> Hold the phone close to your mouth as shown in the picture below. <div style="display: flex; justify-content: space-around; align-items: flex-end; margin: 10px 0;"> <div style="text-align: center;">  <p>NO Too low, too far away</p> </div> <div style="text-align: center;">  <p>YES In front of mouth</p> </div> <div style="text-align: center;">  <p>YES A good distance</p> </div> </div> <ul style="list-style-type: none"> Answer all questions smoothly and naturally in a clear, steady voice. If you don't know the proper way to respond to a test item, you can remain silent or say, "I don't know." Do not take notes or write during the test. When you hear, "Thank you for completing the test", you may hang up. If you wish, you may answer the optional questions at the end of the test. Your personal information will be kept anonymous.

PEARSON

© 2011 Pearson Education, Inc. or its affiliate(s). All rights reserved. Ordinate and Versant are trademarks, in the U.S. and/or other countries, of Pearson Education, Inc. or its affiliate(s).

Side 2 of the Test Paper: Individualized test form (unique for each candidate) showing Test Identification Number, Part A: sentences to read, and examples for all sections.



VERSANT FRENCH TEST

REMINDER: The test begins when you say your name. If you hang up before you complete the test, the test cannot be graded. You cannot reuse the Test Identification Number.



Call: 1-415-738-3800

Test Identification Number

1234 5678

Expires: January 1, 2012

Thank you for calling the Versant testing system.

Please enter your Test Identification Number on the telephone keypad.

Now, please say your name. Now, please say the city and country you are calling from.

Now, please follow the instructions for Parts A through F.

PART	TASK	TEST DETAILS
A	Reading	<p><i>Please read the sentences as you are instructed.</i></p> <ol style="list-style-type: none"> Jean se lève à six heures du matin tous les jours. Quand il fait beau, il va se promener dans le parc pendant une heure. S'il pleut, il va à la salle de gym pour faire un peu d'exercice. Il quitte sa maison pour aller travailler vers huit heures et demie. Mes amis Marc et Michèle attendent un enfant pour le mois de décembre. Ils ont déjà deux petits garçons adorables. Cette fois-ci, ils espèrent que ce sera une petite fille. Ils lui ont déjà choisi un joli prénom. Caroline va obtenir son diplôme au mois de juin. Elle a déjà trouvé du travail. Elle va travailler dans une agence de voyages. Le salaire n'est pas très élevé, mais elle espère obtenir une promotion dans six mois.
B	Repeat	<p><i>Please repeat each sentence that you hear.</i></p> <p>Example: a voice says, "Je pars demain." and you say, "Je pars demain."</p>
C	Questions	<p><i>Now, please just give a simple answer to the questions.</i></p> <p>Example: a voice says, "Où achète-t-on du pain: à la boulangerie ou à la parfumerie ?" and you say, "boulangerie" or "à la boulangerie" .</p>
D	Sentence Builds	<p><i>Now, please rearrange the word groups into a sentence.</i></p> <p>Example: a voice says, "les murs" ... "il a peint" ... "de la cuisine" and you say, "Il a peint les murs de la cuisine."</p>
E	Story Retelling	<p><i>You will hear three brief stories. Each story will be spoken once, followed by a beep. When you hear the beep, you will have 30 seconds to retell the story in French. Try to retell as much of the story as you can, including the situation, characters, actions, and ending. You will hear another beep at the end of the 30 seconds.</i></p>
F	Open Questions	<p><i>You will hear two questions about family life or personal choices. Each question will be spoken twice, followed by a beep. When you hear the beep, you will have 40 seconds to answer the question. You will hear another beep at the end of the 40 seconds.</i></p>

Thank you for completing the test.

Versant French Test - 76 - 11111 - 1



© 2012 Pearson Education, Inc. or its affiliate(s). All rights reserved. Ordinate and Versant are trademarks, in the U.S. and/or other countries, of Pearson Education, Inc. or its affiliate(s).

10. Appendix B: CEFR Mapping

Table 6. General Level Descriptors of the Council of Europe Aligned with Versant French Test Scores.

Level		Council of Europe, 2001 Descriptor	Versant French Test Score
Proficient User	C2	Can understand with ease virtually everything heard or read. Can summarize information from different spoken and written sources, reconstructing arguments and accounts in coherent presentation. Can express him/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in more complex situations.	80 79
	C1	Can understand a wide range of demanding, longer texts, and recognize implicit meaning. Can express him/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibility and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organizational patterns, connectors and cohesive devices.	78 69
Independent User	B2	Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialization. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.	68 58
	B1	Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst traveling in an area where the language is spoken. Can produce simple connected text on topics which are familiar or of personal interest. Can describe experiences and events, dreams, hopes and ambitions and briefly give reasons and explanations for opinions and plans.	57 47
Basic User	A2	Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g., very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.	46 36
	A1	Can understand and use familiar everyday expressions and	35

		very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.	26
	Below A1	Extremely limited proficiency.	25 20

Table 7. Relation of **Versant French Test** Overall scores to Oral Interaction Descriptors based on Council of Europe (2001) framework.

Versant French Test		Oral Interaction Descriptors Based on Council of Europe (2001)
80 79	C2	Conveys finer shades of meaning precisely and naturally. Can express him/herself spontaneously at length with a natural colloquial flow. Consistent grammatical and phonological control of a wide range of complex language, including appropriate use of connectors and other cohesive devices.
78 69	C1	Shows fluent, spontaneous expression in clear, well-structured speech. Can express him/herself fluently and spontaneously, almost effortlessly, with a smooth flow of language. Clear, natural pronunciation. Can vary intonation and stress for emphasis. High degree of accuracy; errors are rare. Controlled use of connectors and cohesive devices.
68 58	B2	Relates information and points of view clearly and without noticeable strain. Can produce stretches of language with a fairly even tempo; few noticeably long pauses. Clear pronunciation and intonation. Does not make errors that cause misunderstanding. Clear, coherent, linked discourse, though there may be some “jumpiness.”
57 47	B1	Relates comprehensibly main points he/she wants to make on familiar matters. Can keep going comprehensibly, even though pausing for grammatical and lexical planning and repair may be very evident. Pronunciation is intelligible even if a foreign accent is sometimes evident and occasional mispronunciations occur. Reasonably accurate use of main repertoire associated with more predictable situations. Can link discrete, simple elements into a connected sequence.
46 36	A2	Relates basic information on, e.g., work, background, family, free time, etc. Can make him/herself understood in very short utterances, even though pauses, false starts, and reformulation are very evident. Pronunciation is generally clear enough to be understood despite a noticeable foreign accent. Uses some simple structures correctly, but still systematically makes basic mistakes. Can link groups of words with simple connectors like “and,” “but,” and “because.”
35 26	A1	Makes simple statements on personal details and very familiar topics. Can manage very short, isolated, mainly prepackaged utterances. Much pausing to search for expressions to articulate less familiar words. Pronunciation is very foreign.
25 20	<A1	Candidate performs below level defined as A1.



About Us

The Knowledge Technologies group of Pearson creates unique technology for automated assessment of speech and text used in a variety of industry leading products and services. These include the Versant line of automated spoken language tests built on Ordinate technology, and WriteToLearn™ automated written summary and essay evaluations using the Knowledge Analysis Technologies™ (KAT) engine.

The Knowledge Technologies group is part of Pearson, the international media company, whose businesses also include the Financial Times Group and the Penguin Group.

Pearson

299 S. California Avenue
Suite 300
Palo Alto, California 94306
USA

4940 Pearl East Circle
Suite 200
Boulder Colorado 80301
USA



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

Pearson now includes Ordinate products and services.

© 2011 Pearson Education, Inc. or its affiliate(s). All rights reserved. Ordinate and Versant are trademarks, in the U.S. and/or other countries, of Pearson Education, Inc. or its affiliate(s). Other names may be the trademarks of their respective owners.

Version 1211B