

Creation of Language Tests for Heritage Learners of the Labrador Dialect of Inuktitut

Marina Sherkina-Lieber

This article pursues several goals related to testing of indigenous language learners: to explain the need for formal testing, to describe the test preparation process, and to present the pilot results of using these tests.¹ I discuss the creation of vocabulary and grammar tests for the purposes of screening participants and assessing the effectiveness of language learning programs for the Labrador dialect of Inuktitut (also called Inuktitut). The speakers of this dialect call it Inuttitut or Inuttut (the Labrador pronunciation of 'Inuktitut' and 'Inuktit', respectively).

The Goal of Testing

Unlike with languages that have thousands, or even millions of speakers and learners, and a large number of language tests, indigenous languages often do not have tests at all, or have a much smaller number. For Labrador Inuttut, tests exist within school language classes and the Labrador Inuktitut Training program (LITP). However, tests for teaching differ from tests for screening: while the former test the knowledge of specific topics learned in class, the aim of the latter is to test overall language knowledge. Therefore, different tests were needed to assess overall proficiency before and after participation in a Master-Apprentice Program.

Testing before the start of a language-learning program is important, first, in order to see whether there is a good fit between the participants and the program, and, second, to determine their proficiency level which, where applicable, ensures correct placement. Testing after the program completion and comparing the test results to the pre-program ones allows us to see how much the program participants learned. Of special interest for testing are non-fluent heritage speakers and receptive bilinguals, who understand much in Inuttitut but do not speak it, or speak very little. Since they already have some language knowledge, it is important to test their knowledge before and after a program to see the program's effect.

The tests discussed in this article were developed for a six-month Labrador Inuttitut Master-Apprentice Program (October 2017-March 2018) at the request of the program organizers, and were used for pre- and post-program testing of the language learners (apprentices). In addition to these tests, learners' participation in a conversation was also evaluated (which is a more traditional approach for testing within a Master-Apprentice Program). Data from the tests and from this speaking assessment was compared.

Labrador Inuttut

Labrador Inuttut, or Labrador Inuttitut, is a dialect of Inuktitut (or Inuktitut) spoken in Inuit communities of Nunatsiavut, the Inuit area of Northern Labrador. Inuktitut, or Inuktitut, is an Eskimo-Aleut language spoken throughout Alaska and Arctic Canada. Both variants are used for both this language in general and its Labrador dialect. (According to Alana Johns' 2009 blog post, Labrador Inuit tend to refer to their dialect as Inuttut in the dialect itself, and as Inuttitut when speaking English.)

While its dialects spoken in Nunavut and Nunavik (Arctic Quebec) are thriving, the Labrador dialect is endangered: most fluent speakers belong to older generations, even though the language is taught at schools, and a large number

Sustaining Indigenous Languages

of Labrador Inuit has only partial knowledge of Inuttitut (Andersen and Johns 2005). Various measures are being taken to preserve and revitalize Inuttitut in Labrador, including language teaching programs. One of these programs is the Master-Apprentice Program.

The Master-Apprentice Program

The Master-Apprentice Program (MAP; also called the Mentor-Apprentice Program) has been developed initially for indigenous languages of California (Hinton 2001), and is now used for a wide variety of languages. The essence of the program is as follows: the apprentice learns the language in one-on-one immersion sessions with the master, who is a fluent speaker of the language in question.

For best results, the organizers decided to recruit apprentices with some pre-existing knowledge of Inuttitut. The rationale was that learners who already have some knowledge would be closer to fluency as a result of the program participation, in comparison to novice learners. Because of this pre-existing knowledge, testing was necessary.

The creators of the Master-Apprentice program suggest oral presentations or conversations as a means of assessing apprentices' progress and generally recommend against formal tests, as outlined, for example, in British Columbia's MAP handbook (Virtue, Gessner, and Daniels 2012). However, as it will become evident in this paper, limiting assessment to speaking skills results in underestimation of apprentices' accomplishments in the program. The Labrador program organizers were interested in tests that would provide more information on the program effectiveness. Even though this is not typically done in MAPs, we decided to try it.

The expected learning outcomes of the program include, first of all, improvement in speaking skills and turning passive knowledge into active, but also increase in the vocabulary size (especially active) and improvement of listening comprehension. Because lack of grammar knowledge is one of the reasons for understanding but not speaking (Sherkina-Lieber 2011; Sherkina-Lieber, Perez-Leroux, and Johns 2011), grammar knowledge also needs to be developed during the program. However, in a MAP, grammar is not taught explicitly, unlike in second or foreign language classrooms. Rather, grammatical rules and functions of grammatical morphemes (such as case or possessive suffixes, for example) are expected to be acquired as in child language acquisition—that is, deduced from input.

Heritage Speakers and Receptive Bilinguals

Many people in Labrador Inuit communities have some knowledge of Inuktitut, though they are far from being fluent in it (Andersen and Johns 2005). This knowledge mostly results in some ability to understand Inuktitut. Most of them were exposed to Inuktitut through their family, starting from their early childhood. They, therefore, qualify as heritage speakers - people who have been exposed to a minority language since a young age and who are to some extent bilingual in the minority and the majority language (definition adapted from Valdés 2000). The term 'heritage speakers' typically refers to immigrant children with respect to their parents' language, but it is also applicable to people raised in indigenous communities where language shift has occurred. As a result of a language shift, children growing in such communities are exposed to the indigenous language much less than to the majority language (the official language of the country). In such a situation, children become fluent in the majority language (the official

language of the country, the language spoken by most people, and also typically the medium of school instruction), while the parents' language is acquired to some extent, often not to full fluency, because of limited exposure to it. This is the situation in Labrador Inuit communities. Many low-proficiency heritage speakers claim to understand Inuttut but not speak it (Andersen and Johns 2005); they are termed receptive bilinguals.

Low-proficiency heritage speakers typically have the following characteristics (see Benmamoun, Montrul and Polinsky 2013 for a review): their understanding is much better than speaking; they have a basic vocabulary (though limited, and their passive vocabulary is larger than active); they also have some basic grammatical knowledge, but much of it is incomplete or inconsistent; if they speak, their speech is slow and effortful, with many long pauses and errors; they often fail to recall a word or build a sentence, and switch to English in such cases.

The main reasons for lower speaking skills in comparison to understanding skills, according to my previous research (Sherkina-Lieber 2011; 2015; Sherkina-Lieber et al. 2011), include gaps in vocabulary and grammatical knowledge, and difficulty recalling a particular word, affix, or structure during speaking. However, their understanding is not perfect either, because they miss some grammatical information, as shown in numerous studies on heritage speakers of various languages (e.g. Montrul 2002; Polinsky 2011, *inter alia*), including Labrador Inuttut (Sherkina-Lieber 2015).

There is a wide variation between heritage speakers in terms of their proficiency level. If program participants are heritage speakers who already have some knowledge of the language in question, it is, therefore, very important to know what they already know and what knowledge they lack before the start of the program. Vocabulary and grammar knowledge need to be tested, and the asymmetry between speaking and understanding skills needs to be taken into account.

Challenges and Ways To Overcome Them

Lack of Literacy

One challenge for testing is that many heritage speakers and receptive bilinguals cannot read and write in Inuttut. In addition, the Master-Apprentice program is traditionally oral, conversation-based and does not include reading and writing instruction. Therefore, its learning outcome is the development of oral skills, namely comprehension and production of spoken Inuttut, but literacy development is not included. This rules out any tests and test formats that involve reading or writing – both pen-and-paper tests and any computerized tests, both pre-program and post-program.

The solution is simple: all tests must be oral. Tasks should involve listening and speaking, but not reading and writing.

Production-Comprehension Asymmetry

For all speakers, speech comprehension is easier than speech production, but this asymmetry is much larger for less fluent speakers, as shown for various heritage languages (Benmamoun et al. 2013; Polinsky and Kagan 2007, among others). Non-fluent heritage speakers are especially likely to have much better speech comprehension than speech production. Therefore, this asymmetry should be taken into account in testing. Oral speech production tests in the form of presentations or conversations give information about the learner's level of speaking proficiency—the central skill that programs such as MAP are expected to develop. However, receptive tests may reveal more knowledge than tests of

Sustaining Indigenous Languages

speech production. Use of only speaking tests (including interviews) will underestimate the heritage speakers' knowledge.

The solution to this problem is to test both production and comprehension, and supplement speaking assessments with formal tests. It is especially important to focus on testing specific aspects of comprehension—not only vocabulary, but also various frequently occurring grammatical structures. While data collection for the analysis of speaking proficiency is obvious—simply recording and analyzing what the learner says,—ways of data collection on understanding are less so. Evaluation of comprehension requires probing whether the message was received with the correct meaning and in all detail. Techniques include comprehension questions, choosing the picture that corresponds to the sentence, acting the sentence out with props, acceptability/grammaticality judgments (“Is this a good sentence?”), and others. The drawback of comprehension tests is that they are less “natural” than a conversation or an interview. On the other hand, there is an additional benefit of formal tests of comprehension: their results provide information about the knowledge of specific grammatical variables. When examining free-form speaking assessments such as presentation, conversations and interviews, it may be impossible to come to conclusions regarding knowledge (or lack thereof) of at least some of linguistic variables because learners can avoid constructions that they have not fully mastered.

The Polysynthetic Nature of the Language

Another challenge is provided by the language itself. Inuttut is a polysynthetic language, with rich morphology. Vocabulary knowledge, especially in a polysynthetic language, is mostly knowledge of the roots. The problem for testing is that roots in a polysynthetic language (especially verb roots) are not always possible to present in isolation (Sherkina-Lieber and Helms-Park 2015). A word in Inuttut can be relatively short, as in (1), or long, consisting of many morphemes, and equivalent to a whole sentence in English, as in (2).

- (1) *pisujunga*
pisu-junga
walk-DECL.1SG²
'I am walking'

- (2) *nipalaattailiKujaunginnalaukKugut*
nipalaa-ttaili-Ku-jau-nginna-lauk-Kugut
make.noise-prevent-tell-PASS-always-DIST.PAST-DECL.1PL
'We were always told to be quiet'

While nominal roots can occur in isolation because the absolutive case marker is phonologically null for singular non-possessed nouns, verbal roots never do; they have at least an agreement marker, but often many other affixes (tense, aspect, negation, etc.). Affixes can be stacked, as in (2). A word can contain as many as ten morphemes (Dorais, 1988). Some Inuttut verbs (including equivalents of be, have, get) obligatory incorporate nouns or other verbs (e.g. *-Ku-* 'tell', *-guma-* 'want'). Some adjective-like elements are suffixes. This rich morphology makes vocabulary testing difficult.

The solution is that for vocabulary tests, decisions must be made on how to present lexical items that never occur on their own. Specific forms must be chosen to act as default forms for each particular type of words. These forms must be morphologically simplest, with a minimal amount of affixes. The choice of the forms for the Labrador Inuttut vocabulary test is discussed in the corresponding section below.

Testing As Many Linguistic Variables As Possible vs. Limits on Test Length

For a test to be representative of an individual's general knowledge of Inuttitut, many linguistic variables need to be included, and, ideally, several instances of each variable. However, such a test would not be feasible because it would take too much time. Therefore, representative and/or critical variables must be chosen.

The solutions for vocabulary and grammar tests are as follows. For vocabulary testing, the choice of items must reflect the learning process. That is, tests should include those words that are likely to be learned during the program. Still, it is impossible to include all such items in a test. Therefore, representative items can be chosen: a small number of items from each of the semantic fields likely to be covered in the program (such as, for example, family members (e.g. *panik* 'daughter'), food and cooking (e.g. *mannik* 'egg'), body parts (e.g. *ijik* 'eye'), feelings (e.g. *Kuviasutuk* 'he/she is happy'), and others). Items should also span a range of frequency of use, that is, include lexical items of high, medium and low frequency.

For grammar testing, variables chosen should all be frequently used and necessary to produce common everyday sentences (such as subject-verb agreement, plural, possessive, cases, tenses, and so on). However, they should vary in complexity and vulnerability in heritage language. Benmamoun et al. (2013), after reviewing a number of heritage language studies, concluded that (low-proficiency) heritage speakers tend to have a better knowledge of subject-verb agreement morphemes than case morphemes, and better knowledge of tense than aspect.

Because of the length constraint on the number of test items, the longer the list of grammatical variables, the fewer items per variable it would be possible to include. Therefore, for a comprehensive test, a small number of items per grammatical variable can be included.

Linguists Are Not Fluent Native Speakers and Fluent Native Speakers Are Not Linguists

The ideal situation for a test preparation is when it is done by linguistic professionals who are also fluent in the language. However, this is not always possible. While it is necessary for the linguists who create tests to have a good understanding of the structure of the language in question, they often are not fluent native speakers of the language. On the other hand, fluent native speakers in the community often have no background in linguistics and/or language testing.

The solution is that, for best results, linguists who have substantial knowledge of the language but are not fluent native speakers themselves should collaborate with fluent speakers (who are not linguists). This was accomplished in our case via linguistic fieldwork in three steps presented below:

(3) Test creation in consultation with fluent native speakers

STEP 1: The linguist with knowledge of the language in question chooses linguistic variables to test.

STEP 2: The linguist creates test words and sentences with the help of 1-3 fluent speakers.

STEP 3: The test words and sentences are then checked by 2-3 other fluent speakers for grammatical well-formedness, felicity, word frequency, and cultural appropriateness.

Step 3 can also work as a pretest for the potential test items (words and/or sentences). If the additional two or three fluent speakers do the same task that the learners will do during the test and give the answers that are expected (on the basis of work done in Step 2), the items are suitable to be used for testing.

Sustaining Indigenous Languages

Otherwise, the test preparation process should go back to Step 2 and create new items instead of the problematic ones.

No Linguists Available To Conduct the Test

As noted above, a community often would not have individuals with linguistic or language testing background. While the development of tests can be a one-time project, the long-term use of tests for a program requires regular availability of individuals who can conduct the tests (and funding to cover their travel costs if they are from outside of the community).

The solution is that the tests should be simple enough to be administered by fluent speakers who might have no background in linguistics, so that the community could make use of these tests without the need to request outside help.

The linguist who develops the tests, therefore, has to ensure that the instructions for test administration should be easily understandable by a non-linguist, and therefore should not contain linguistic terminology unknown to non-linguists. Scoring should also be as straightforward as possible, and so should be interpretation of the results. In brief, the test administrator simply needs to mark the answers as correct or incorrect (with an answer key provided), count the number of correct answers, and then calculate the percentage of correct answers.

To achieve this, a document was created for each test in our program. The document included the following: the instructions, the list of test items, the correct answers for each item and a way to mark the answer (such as a box for a check mark).

The instructions contained an explanation of how to present the test items and how to explain the task to the learner. Since the tests were oral, presentation of test items (words or sentences) typically included reading them out loud by the person administering the test; some tests also involved showing pictures (see below for the instructions for specific tests).

The list of test items was given in the order in which they were to be presented to the learner. For all tests, the items were listed in the pseudo-randomized order so that there is not too many similar sentences in a row.

The correct answer for each test item was printed next to the item. The person who administered the test needed to put a check mark in the box next to the item if the answer was correct, and a cross mark if it was not. If there were more than one possible correct answers, all were listed, and the tester was instructed to accept any of them. If an answer was not listed, but the tester considered it correct, they were instructed to write it down on the scoring sheet and mark it as correct.

If the test task was to choose between two or more alternatives, all the alternatives were listed, and the correct alternative was marked. For such tests, the person administering the test needed to circle the answer chosen by the learner and mark it as correct or incorrect with a check or a cross mark, respectively.

Of course, training is still required for the non-linguists who will administer the test. In the MAP program, the masters were trained by the author to administer the tests to their own apprentices. When the master for a particular apprentice was not available, another master could administer the tests. I provided training as part of the MAP orientation. For each test, we went through the testing package with the masters. Then masters did the pre-program testing with their apprentices in my presence so that I could answer any questions, help with scoring, and/or correct any issues in the testing process.

In the next section, I discuss the specific tests that I developed for MAP in consultation with Labrador Inuttut fluent speakers.

**The Tests of Vocabulary and Grammar
in the Labrador Inuttut Master-Apprentice Program**

As explained before, pre- and post-program testing should include both speech production and speech comprehension. The formal tests included both speech production and speech comprehension in Inuttut, and examined the knowledge of vocabulary and grammar. In addition to the formal tests, speech production was tested in the non-formal way suggested by the creators of the Master-Apprentice Program: each apprentice had a choice of preparing a presentation on any topic or participating in a conversation/interview with two fluent speakers other than his/her master.

The Vocabulary Test

Lexical knowledge of the apprentices was tested in both production and comprehension. Therefore, the vocabulary test consisted of two parts. In one part, learners translated selected words from English to Inuttut; in the other part, they translated the same words in the opposite direction. Translation from Inuttut to English, which was the passive vocabulary test, can be regarded as a comprehension test, since the learners had to understand Inuttut words and find English equivalents for their meaning. Translation from English to Inuttut—the active vocabulary test—can be then regarded as a speech production test, since the learners had to say the Inuttut word corresponding to the meaning expressed by an English word.

As Inuttut is polysynthetic, the vocabulary test was focused on the knowledge of roots. Therefore, words with the minimum of non-root morphemes were chosen. The word types used in the test, with the number of items per type, are listed in (4). The same words were used in the pre- and post-program test versions. Nouns and noun-like adjectives were presented in the absolutive case singular form, the suffix for which is phonologically null (e.g. *siutik-Ø* ‘ear’). As there are no verb forms without suffixes or with phonologically null suffixes, verbs were presented with only one portmanteau suffix expressing third person singular, antipassive, participial mood form (e.g. *nigi-juk* ‘he/she is eating’, where the root is *nigi-* ‘eat’; *-juk* is the suffix). The participial mood was selected because in the Labrador dialect, it is used in third person forms instead of declarative in default situations; the use of declarative in third person is reserved for conveying special meanings (Johns 1995). The antipassive, rather than ergative, was chosen because, though Inuktitut in general is an ergative language in which verbs bear suffixes that reflect agreement with both the subject and the object, the Labrador dialect mostly uses verbs in the antipassive form, especially in isolated sentences which are not part of a connected discourse (Johns 1999, 2001).

Words that are listed as common expressions are multiword expressions only in the English versions; in Inuttut, they are single words.

- (4) Word types selected for the vocabulary test
- 45 nouns
 - 5 noun-like adjectives
 - 45 verbs
 - 9 common “expressions” (greetings; “yes”, “no”, “thank you”, “nice weather”, “I don’t know”, etc.)

The test procedure was as follows. The active vocabulary was tested first. The learner heard an English word read by the person administering the test, then translated it orally into Inuttut. The test administrator marked the word with a check mark if it was translated correctly, and with an X if the translation was

Sustaining Indigenous Languages

incorrect, or the learner was unable to translate. After that, the next word was presented, and so on, until all the words in the list were translated. The passive vocabulary was tested after. When time is limited, it is recommended to present for Inuttut-to-English translation only those words which the learner failed to translate from English to Inuttut (because active knowledge implies passive knowledge, but not the reverse).

The results are the percentages of correctly translated words (separately for passive and active vocabulary subtests). Heritage speakers and receptive bilinguals who translate 70-80% of words correctly are likely to have relatively good comprehension skills (Sherkina-Lieber 2011).

Grammar Tests

The grammar test consisted of three parts, or subtests. One part, a cloze test with pictures, tested grammar knowledge in speaking (that is, it was a speech production test). The second and third parts tested receptive grammar knowledge. The second part focused on grammatical morphemes' contribution to meaning (e.g. past tense, habitual aspect, etc.), and the last part focused on well-formedness of sentences. Within each subtest, the same linguistic variables, but different sentences, were used in pre- vs. post-program tests. In each subtest, the percentage of correct answers was calculated.

Productive Grammar: A Cloze Test

The goal of this subtest is to test the use of correct grammar in speech production. When we produce even a simple sentence, we need to produce each word in its required form, with all the necessary functional morphemes, such as, for example, case and number suffixes for nouns and agreement suffixes for verbs (Inuktitut does not have prefixes; all the functional morphemes in it are suffixes). Unlike in a free-form speaking assessment, this test was designed to elicit specific words in specific forms.

Eight linguistic variables were selected for this subtest. Two items per variable were created, for the total of 16 items. The linguistic variables selected for this subtest are listed in (5).

(5) Linguistic variables selected for the cloze test

- Oblique case on nouns
- Plural suffix on nouns
- Possessive suffix on nouns
- Verb agreement suffix: third person singular, participial mood
- Verb agreement suffix: third person dual, participial mood
- Verb agreement suffix: third person plural, participial mood
- The conjunction "and" (as a suffix -lu or a free word amma(lu))
- Noun incorporation

The agreement suffixes in this subtest were for antipassive verbs only, reflecting subject-verb agreement (but not subject-object verb agreement).

Each test item was a sentence with the last word containing one of the test variables. Each sentence was matched with a photo reflecting what was described in the sentence. The procedure was as follows. The person administering the test shows the learner a photo and reads the corresponding sentence except the last word. The learner's task is to look at the photo and finish the sentence by producing the last word. Examples are shown in (6) for case and (7) for plural; the word in brackets is the one that the learner is expected to produce.

- (6) (PHOTO: A man is looking at a house)
Angutik taku-juk... (illu-mik)
man.ABS see-PART.3SG. **house-oblique.case**
'The man sees... (a house)'
- (7) (PHOTO: Four dogs)
Ukua sitamat... (Kimmi-it)
These four **dog-PL**
'Here are four... (dogs)'

The person administering the test puts a check mark next to a sentence only if the word produced by the learner has the correct form. In (6), it is not enough to produce the word for 'house'; it must have the case suffix *-mik*. In (7), the word for 'dog' has to have the plural suffix *-it*. If the required suffix is missing, or replaced by an incorrect suffix, or if the learner fails to produce the word, the item is marked by an X. The results of this subtest are the percentage of correctly produced suffixes.

Receptive Grammar: Comprehension Questions

This subtest probes knowledge of the contribution of functional morphemes (suffixes) to meaning. Heritage speakers and receptive bilinguals are aware that, while some of them may understand much of what they hear, they do not understand everything. One of the most important factors in this is lack of grammar knowledge and lack of knowledge about the contribution of suffixes to meaning. Without this knowledge, it is impossible to distinguish, for example, between past events and events that will happen later, or between passive and active, and so on. For this subtest, the functional morphemes that provide easily recognizable contribution to meaning were selected, and matched in contrasting pairs; they are listed in (8). Five contrasts were selected, and four items were created for each, for the total of 20 items. The items were sentences that contained a word with a given functional morpheme. Each sentence was accompanied by a comprehension question about the contribution of that functional morpheme. The sentences and the questions were created so that the question could only be answered correctly if the learner recognizes the functional morpheme; other words contained no clues to the answer. Within each contrast, two items contained one morpheme, and the remaining two, the other (e.g. in the Tense: Past vs. Future contrast, two items contained a verb with a past tense suffix, while two other items contained a verb with a future tense suffix.)

- (8) *Grammatical contrasts selected for functional morpheme comprehension test*
- Case: allative ("to") vs. ablative ("from")
 - Voice: passive vs. active
 - Tense: past vs. future
 - Remoteness in tense: near past/future vs. distant past/future
 - Aspect: habitual vs. ongoing (2 items), perfective vs. progressive (2 items)

The procedure was as follows. The person who administered the test read the sentence, and then read the corresponding comprehension question. The learner's task was to listen to the sentence and answer the comprehension questions (in English). The comprehension questions were of the forced-choice type. They are asked in English, and were expected to be answered in English, in order to maximize performance and not to contaminate the answer by the necessity to understand the question in addition to understanding the sentence, and to speak

Sustaining Indigenous Languages

Inuttitut. In (9), the contrast between the allative and the ablative case and the contribution of the allative case suffix is tested (the allative suffix *-mut* contributes the meaning similar to that of the English preposition *to*). In (10), the past tense and future tense contrast and the contribution of the distant future suffix *-lât-* is tested.

- (9) Mary *appali-juk* *ânniasiuvi-mmut.*
Mary run-PART.3SG hospital-ALLATIVE
'Mary is running **to** the hospital'
QUESTION: Is Mary running to the hospital or from the hospital?

- (10) Sally *tuttu-vini-mmik* *nigi-lât-tuk.*
Sally caribou-former-OBLIQUE eat-DIST.FUTURE-PART.3SG
'Sally **will** eat caribou meat'
QUESTION: Will Sally eat caribou meat later, or did she already eat caribou meat?

If the learner chooses the correct answer from the two options contained in the question, the person administering the test puts a checkmark next to that item; otherwise, he/she puts an X there. The percentage of correct answers is then calculated.

Receptive Grammar/Metalinguistic: Grammaticality Judgments

Some grammatical morphemes do not give a clear, testable contribution to meaning. Rather, they are necessary for well-formedness of the sentence. Such morphemes cannot be tested with the help of comprehension questions. The goal of this last grammatical subtest is to probe the knowledge of grammatical morphemes needed for well-formedness. Their knowledge is tested via the learners' ability to detect violations of well-formedness, known as grammatical errors.

The items in this task are pairs of sentences in which one is a well-formed sentence, and the other differs from it only in one morpheme, such that the morpheme is missing, incorrect, or in an incorrect position. As the result of the difference, the latter sentence is ungrammatical. The list of grammatical errors selected for this subtest is presented in (11). There are six error types, with three items per error; therefore, there are 18 sentence pairs. The order of grammatical and ungrammatical sentences is counterbalanced: in half of the pairs, the grammatical sentence comes first, while in the other half, the ungrammatical sentence comes first.

- (11) Types of grammatical errors for grammaticality judgments
- Missing agreement suffix on a verb
 - Subject-verb number mismatch
 - Declarative mood instead of interrogative
 - Incorrect order of tense and negation suffixes
 - Missing an oblique case suffix
 - An incorporated noun moved

In this subtest, the person administering it reads a pair of sentences. The learner has to tell which sentence is better, or point out which sentence contains a grammatical error. The examples are provided in (12), where the ungrammatical sentence (marked with an asterisk) contains the singular verb agreement suffix instead of the plural one, and (13), where the incorporated noun, which must occupy the position before the incorporating verb, is moved to the postverbal position in the ungrammatical sentence.

Creation of Language Tests for Heritage Learners

- (12) a. *Sugusi-it* *nigi-jut.*
 Child-PL eat-PART.3PL
 ‘Children **are eating**’
- b. **Sugusi-it* *nigi-juk.*
 Child-PL eat-PART.3PL
 ‘*Children **is eating**’
- (13) a. *Annak* *nasa-liu-juk.*
 Woman hat-make-PART.3SG
 ‘The woman is making a hat/hats’(lit. ‘The woman is **hat-making**’)
- b. **Annak* *liu-juk* *nasak.*
 Woman make-PART.3SG. hat

The person administering the test marks the answer as correct if the learner correctly points out which sentence is grammatical (“good,” “well-formed,” “has no errors”) and/or which one is ungrammatical (“bad,” “incorrect,” “contains an error,” “broken Inuttitut”). If the learner accepts the ungrammatical sentence as a good one and rejects the grammatical sentence as a bad one, or if he/she states that both sentences are good, or both are bad, the answer is recorded as incorrect. The percentage of correct answers is then calculated.

Test Results in the Master-Apprentice Program

The tests described above were administered to the apprentices in the Labrador Inuttut MAP before and after the program. As this was a pilot program, only five master-apprentice pairs participated, one from each Labrador Inuit community. Only three pairs stayed in the program till the end, and only two apprentices could be tested both pre- and post-program. The post-program testing of both apprentices and the pre-program testing of Apprentice B were done in my presence.

The vocabulary test results are presented in Table 1. The results showed the following. First, the apprentices started the program already with some vocabulary knowledge; their passive vocabulary was slightly larger than active. This is not unexpected for the heritage language speakers and learners. This asymmetry disappeared post-program for Apprentice B, but not for Apprentice A. Second, some vocabulary growth during the program was found, especially for active vocabulary, and especially for Apprentice B. On the other hand, during the post-program testing, apprentices were sometimes unable to translate the words that they had successfully translated during the before-program testing. Therefore, MAP helped the participants to increase their active vocabulary size.

Table 1. The vocabulary test results. Percentage of correct answers (and raw numbers).

	Apprentice A		Apprentice B	
	before	after	before	after
Active vocabulary (English to Inuttut)	84.3% (91/108)	86.1% (93/108)	83.3% (90/108)	99.1% (107/108)
Passive vocabulary (Inuttut to English)	93.5% (101/108)	93.5% (101/108)	88.9% (96/108)	99.1% (107/108)

The results of the grammar tests showed that the apprentices started the program with some basic grammar knowledge, but also inconsistent performance and lack of knowledge of some aspects of grammar. This is also typical of non-fluent

Sustaining Indigenous Languages

heritage speakers (Sherkina-Lieber 2011). The results of the grammar tests are shown in Table 2 below.

Signs of grammatical difficulties were evident not only in the pre-program tests, but also in the post-program ones. In the cloze test, which tested grammar in speech production, the apprentices made grammatical errors, such as producing incorrect forms (e.g. singular instead of plural) or omitting functional morphemes (e.g. being unable to produce possessive suffixes). In the comprehension test, the apprentices misunderstood the contribution of at least some functional morphemes (e.g. case markers roughly equivalent to the meaning of 'to' and 'from', as in *ilinniavi-mut* 'to school' vs. *ilinniavi-mit* 'from school') or did not notice the contribution of a functional morpheme (e.g. they did not notice a passive suffix and therefore misinterpreted the sentence as active). In the grammaticality judgment test, the apprentices sometimes incorrectly judged grammatical sentences as ungrammatical and vice versa.

Table 2. The grammar test results. Percentage of correct responses (and raw numbers).

Test	Apprentice A		Apprentice B	
	Before	After	Before	After
Productive grammar: cloze test	100% (16/16)	50% (8/16)	62.5% (10/16)	75% (12/16)
Receptive grammar: comprehension questions	70% (14/20)	30% (6/20)	65% (13/20)	55% (11/20)
Receptive/metalinguistic grammar knowledge: grammaticality judgments	83.3% (15/18)	55.6% (10/18)	77.7% (14/18)	83.3% (15/18)

Unfortunately, little or no grammar improvement during the program was found when the pre-program and post-program test results were compared. In the case of Apprentice A, it appears as if this apprentice lost grammatical knowledge rather than gained it; this can be real, or can be possibly due to testing problems (such as the master not noticing some of the errors in the pre-program test).

Speaking assessment results were mostly consistent with test results. Both apprentices have chosen to participate in a conversation with fluent speakers (other than their masters), as opposed to prepare a speech. The apprentices produced mostly short sentences that contained some grammatical errors and lexical errors, somewhat limited vocabulary, and some failures to produce a word (because of inability to recall it) or finish a sentence. They mostly answered questions rather than asked them or initiated topics.

Notably, Apprentice A, who was barely able to say something in the beginning of the speaking assessment, then had many failures to produce a word or finish a sentence and was very nervous about conversing in Inuttut, showed more knowledge on the tests than in the conversation. This apprentice's vocabulary test results showed a significant amount of vocabulary knowledge, and the grammar test results showed that this apprentice knew some basic aspects of grammar, but there was still much to be learned. These specific grammatical difficulties explain this apprentice's struggle with speech production.

Speaking assessments give an overall picture, but grammar tests give more details on what grammatical properties are known or not known. Speaking assessments also show communicative competence in the language in question; grammar and lexical tests show knowledge even in those who have difficulties participating in a conversation.

As evident from the experience with the apprentice who showed more knowledge on the formal tests than in an informal speaking assessment, being

at least partially successful on formal tests actually give apprentices confidence and reveal that they know more than they thought. On the other hand, confidence is an issue when an apprentice needs to carry out a conversation or a presentation, and if the apprentice fails to say as much as he/she wanted, it leads to loss of confidence.

Giving tests in the middle of a MAP can also be helpful for making the most of the program. Apprentices and masters can learn from the test results about the areas of language knowledge in which the apprentice needs improvement – especially grammatical knowledge, which is necessary to build sentences.

Conclusion

Formal grammar and vocabulary tests are helpful tools for assessing learners' abilities and progress. They complement speaking assessments such as presentations and conversations by giving information on speech comprehension (in addition to speech production) and on knowledge of specific words and specific aspects of grammar. Vocabulary and grammar need to be tested both in comprehension and in production. In low-proficiency heritage speakers and receptive bilinguals, tests of speech comprehension may reveal more knowledge than tests of production. Tests for heritage speakers should be oral, and designed so that they are easy to administer, score and interpret for non-specialists.

Notes

- ¹ I wish to thank for help: Silpa Suarak, Language Coordinator, Nunatsiavut Government; Brenda Jararuse, Director of Culture, Nunatsiavut Government; Inuttitut language consultants; and, last but not least, the participants in the Labrador Inuttitut Master-Apprentice program. The tests described in this article were developed for the 2017-2018 Labrador Inuttitut Master Apprentice Program and funded by Traditions and Transitions Research Partnership, Memorial University of Newfoundland.
- ² Abbreviations in the examples: SG – singular; PL – plural; DECL; – declarative mood; PART – participial mood; PASS – passive; DIST.PAST – distant past tense; DIST.FUTURE – distant future tense; OBLIQUE – oblique case; ALLATIVE – allative case.

References

- Andersen, Catharyn, and Alana Johns. 2005. "Labrador Inuttitut: Speaking into the Future." *Études/Inuit/Studies* 29:187-205.
- Benmamoun, Elabbas, Silvina Montrul, and Maria Polinsky. 2013. "Heritage Languages and their Speakers: Opportunities and Challenges for Linguistics." *Theoretical Linguistics* 39, no. 3-4:129-181.
- Dorais, Louis-Jacques. 1988. *Tukilik: An Inuktitut Grammar for All*. Inuit Studies Occasional Papers 2. Quebec: Association Inuksiutiit Katimajit Inc. and Groupes d'Études Inuit et Circumpolaires.
- Hinton, Leanne. 2001. "The Master-Apprentice Language Learning Program." In *The Green Book of Language Revitalization in Practice*, edited by Leanne Hinton and Ken Hale, 217-226. Leiden: Brill.
- Johns, Alana. 1995. "On Some Mood Alternations in Labrador Inuttut." In *Grammatical Relations: Theoretical Approaches to Empirical Questions*, edited by Clifford S. Burgess, Katarzyna Dziwirek, and Donna Gerdts, 131-151. Center for the Study of Language and Information, Stanford University.

Sustaining Indigenous Languages

- Johns, Alana. 1999. "The Decline of Ergativity in Labrador Inuttut." In *Papers from the Workshop on Structure and Constituency in Native American Languages*, MIT Occasional Papers in Linguistics 17, edited by Leora Bar-el, Rose-Marie Dechaine, and Charlotte Reinholtz, 73-90.
- Johns, Alana. 2001. "An Inclination Towards Accusative." *Linguistica Atlantica* 23:127-144.
- Johns, Alana. 2009. "Inuttitut or Inuttut?" A Blog Post on the Labrador Community Grammar Blog. Accessed February 21, 2019. <https://ukausivut.wordpress.com/2009/07/31/inuttitut-or-inuttut/>.
- Montrul, Silvina. 2002. "Incomplete Acquisition and Attrition of Spanish Tense/Aspect Distinction in Adult Bilinguals." *Bilingualism: Language and Cognition* 5:39-68.
- Polinsky, Maria. 2011. "Reanalysis in Adult Heritage Language: New Evidence in Support of Attrition." *Studies in Second Language* 33, no.2:305-328.
- Polinsky, Maria, and Olga Kagan. 2007. "Heritage Languages: In the 'Wild' and in the Classroom." *Language and Linguistics Compass* 1, no. 5:368-395.
- Sherkina-Lieber, Marina. 2011. "Knowledge of Labrador Inuttitut Functional Morphology by Receptive Bilinguals." PhD diss., University of Toronto.
- Sherkina-Lieber, Marina. 2015. "Tense, Aspect, and Agreement in Heritage Labrador Inuttitut: Do Receptive Bilinguals Understand Functional Morphology?" *Linguistic Approaches to Bilingualism* 5, no. 1:30-61.
- Sherkina-Lieber, Marina, and Rena Helms-Park. 2015. "A Prototype of a Receptive Lexical Test for a Polysynthetic Heritage Language: The case of Inuttitut in Labrador." *Language Testing* 32:419-442.
- Sherkina-Lieber, Marina, Ana-Teresa Pérez-Leroux, and Alana Johns. 2011. "Grammar Without Speech Production: The Case of Labrador Inuttitut Heritage Receptive Bilinguals." *Bilingualism: Language and Cognition* 14, no. 3:301-317.
- Valdés, Guadeloupe. 2000. "Teaching Heritage Languages: An Introduction for Slavic-Language-Teaching Professionals." In *Learning and Teaching of Slavic Languages and Cultures: Toward the 21st Century*, edited by Olga Kagan & Benjamin Rifkin, 475-403. Bloomington, IN: Slavica.
- Virtue, Hannah, Suzanne Gessner, and Deanna Daniels (Xway'Waat). 2012. *B.C.'s Master-Apprentice Language Program Handbook*. First Peoples' Cultural Council (FPCC). Accessed February 21, 2019. http://www.fpcc.ca/files/PDF/MAP_HANDBOOK_2012.pdf.