Utrecht University

Language assessment at a distance

The TeleTaalTest-NL (TTT-NL), a Dutch tele-language test for patients with neurological disorders

Thesis RMA Linguistics
Utrecht University

Student name: Naomi A.A.M. Legius
Student number: 6656609
Supervisor Erasmus University MC: Dr. Djaina Satoer
Supervisor Utrecht University: Prof. dr. Sergey Avrutin
Second reader: Dr. Shalom Zuckerman
Date: November 19, 2020

Abstract

Background. Unforeseen situations such as the recent outbreak of COVID-19 negatively impact aphasia assessment and diagnostics in the form of canceled or postponed face-to-face language assessments. A proposed solution is the TeleTaalTest-NL (TTT-NL), a language test that can be administered by telephone, which ensures that patient screening, monitoring, and offering advice still continues at a distance.

Methods. Items and subtests of the existing version of the TTT-NL were revised and assessed in a pilot group consisting of 10 healthy BA and MA students. Subsequently, the test version of the TTT-NL was administered in a group of 28 healthy individuals, and correlated to the Diagnostic Instrument for Mild Aphasia (DIMA), an in-person language test.

Results. Pilot items with an 80% correctness rate or higher were included for the development of the test version. Healthy control groups from both the pilot and the norm study successfully completed the TTT-NL and an effect of age, education, and hearing was found. Only the semantic selection task significantly correlated with in-person language assessment, which is most likely due to the small sample size.

Conclusion. The improvement of the TTT-NL was successful and it is a feasible alternative for face-to-face assessment. However, there is no evidence yet that demonstrates the TTT-NL is a reliable replacement for in-person assessment. Hence, further research of a larger sample size into the relationship between tele-testing and in-person assessment is advised.

Keywords: aphasia, language test, telephone assessment, TeleLanguage Test, neurological disorders

Contents

Chapter 1: Introduction

Each year, approximately 140.000 individuals of the Dutch population acquire brain damage of which 40.000 individuals suffer from permanent damage such as aphasia, a language disorder that affects language perception and/or production across modalities depending on the type of neurological lesion (Bastiaanse, 2011; Hersenz, n.d.). The number of people who have been diagnosed with aphasia increases each year as a consequence of a neurological cause such as stroke, head trauma, a brain tumor, or an inflammation such as meningitis. Research shows that aphasia also frequently is comorbid with neurodegenerative diseases such as Parkinson's, Alzheimer's and Huntington's disease (Bastiaanse, 2011; Hallowell, 2017).

Under regular conditions, aphasic patients are referred to a clinical linguist and/or a speech therapist at the hospital or a rehabilitation center for extensive diagnostics. Several test-batteries / tests have been designed / adapted for cognitive (test-batteries) and aphasia diagnostics (and follow-up after language treatment) in Dutch, such as the Akense Afasie Test (AAT) (Graetz et al., 1992), a test-battery that originally determined the nature and severity of aphasia by means of test scores and now describes the disorder by modality and linguistic level through the assessment of language comprehension, spontaneous speech, verbal repetition, writing-to-dictation, reading aloud and naming, the Comprehensive Aphasia Test-NL (CAT-NL) (Visch-Brink et al., 2014), a test-battery for language comprehension, verbal repetition, naming, reading, writing, and the influence of aphasia on the patient's daily life, the ScreeLing (Visch-Brink et al., 2010), a test-battery that assesses semantics, phonology, syntax, language comprehension, language production, reading, and verbal repetition, the more recently developed Diagnostic Instrument for Mild Aphasia (DIMA) (Satoer et al., 2019), a test-battery that assesses mild impairments at the level of phonology, semantics, syntax, , and spontaneous speech in context, the Boston Naming Test (BNT) (Kaplan et al., 1983), a picture-naming test that assesses word retrieval, and the shortened Token Test (De Renzi & Faglioni, 1978), a test that determines aphasia severity and assesses language comprehension. These tests have been designed for face-to-face assessment and have also been standardized as such in a healthy population. Travel distance in the Netherlands usually is not a hindrance for aphasia diagnostics in a relatively compact country. However, unforeseen situations such as the recent outbreak of the COVID-19 virus have a big impact on in-person language assessments and on the health system in general as all "non-urgent" appointments - including in-person language assessment - are cancelled or postponed.

A proposed solution for the cancelled outpatient appointments is the TeleTaalTest-NL (TTT-NL) (Satoer et al., 2020a, see Van Dijk, 2018; and Ruijs, 2018 for previous versions) a language test that can be administered by telephone. The TTT-NL is based on the English TeleLanguage Test (De Witte et al., 2019), originally designed as a pre- and postoperative test for American patients with a brain tumor without access to a hospital or clinic due to great distances and/or no access to video calling. The subtests of the TeleLanguage Test are based on English aphasia tests, such as the Western Aphasia Battery (WAB) (Kertesz, 2007) and the Boston Diagnostic Aphasia Examination (BDAE) (Goodglass et al., 2000), and items from the Dutch Linguistic Intraoperative Protocol (DuLIP) (De Witte et al., 2015) and the DIMA. Before the development of the TeleLanguage Test, studies have shown that cognitive screening by telephone using tests such as the Telephone Montreal Cognitive Assessment (t-MoCA) (Pendlebury et al., 2013) and the Telephone Interview for Cognitive Status (TICS-m) (Van den Berg et al., 2012) is a feasible, precise, and reliable way to measure cognitive functions in senior citizens and various clinical populations (i.e. stroke, TIA, and Alzheimer's disease) (Castanho et al., 2014; Mitsis et

al., 2010; Pendlebury et al., 2013; Zietemann et al., 2017). Cognitive testing by phone appeared to be successful, however none of the available language tests for face-to-face assessment had been validated for telephone assessment, nor did a tele-language test exist yet. Hence, the TeleLanguage Test was developed. A validation study by De Witte and colleagues (2019) revealed that the TeleLanguage test is a feasible and preliminary valid test battery as an alternative for face-to-face assessment in brain tumor patients as both versions of the test proved to be reliable for the detection of mild aphasia and were similar to, or even more sensitive than, in-person assessment of the quick aphasia battery (QAB) (Wilson et al., 2018). These results indicate that telephone language assessment is a feasible and valid supplement to/alternative for face-to-face assessment during the COVID-19 pandemic and other situations in which in-person language assessment is not an option.

The current norm data of the TTT-NL consists of 28 healthy controls (17 women) between the ages of 40 and 70 years old (Satoer et al., 2020a). Although the current version of the TTT-NL is a preliminary valid tool for language testing, it needs revision as some items did not yield the desired results, such as word finding items to which multiple alternative answers could be given (Van Dijk, 2018). Furthermore, scores of only three patients on a previous version of the test have been correlated to their neuropsychological screening (NPS) scores (Ruijs, 2018). Therefore, the present study aims to improve the items and subtests of the TTT-NL and supplements its norm data with at least 20 healthy individuals for a more reliable reference group. This dataset is then compared to previously collected healthy control data of the DIMA (Satoer et al., 2019), a face-to-face test, to assess the construct validity of the TTT-NL.

1.1 Aphasia

Aphasia is a language disorder that can be acquired as a consequence of stroke (most common, see Hallowell, 2017), head trauma, an inflammation (e.g. meningitis), neurodegenerative diseases such as Alzheimer's, Parkinson's, and Huntington's disease, or brain tumor (Bastiaanse, 2011; Hallowell, 2017). Aphasia should not be confused with speech disorders such as apraxia of speech (i.e. deficit in programming articulatory movement) and dysarthria (i.e. articulatory disorder as a consequence of muscle weakness or the impaired ability to control the muscles involved in speech), where the problem is of a motoric nature (Bastiaanse, 2011; Grossman & Irwin, 2018; Hallowell, 2017). Lesions can occur in both the cerebral cortex and subcortical brain structures, the latter supposedly playing a decisive role on aphasia severity and the degree of recovery (Bastiaanse, 2011; Incekara et al., 2019; Marchina et al., 2011; Sierpowska et al., 2017). Given the fact that in 95 to 98% of the righthanded individuals and 70% of the lefthanded individuals language is represented in the left hemisphere, aphasia is most likely caused by a lesion in this hemisphere. Although rare, there are cases in which aphasia in righthanded individuals is caused by a lesion in the right hemisphere (i.e. crossed aphasia) (Bastiaanse, 2011).

Depending on the type and location of the lesion, the patient experiences hindrance in one or more linguistic levels. The main levels include phonology, semantics, and syntax and they can be affected in different modalities (i.e. language production and/or language comprehension in speaking, writing or reading) (Bastiaanse, 2011; Satoer et al., 2018). Based on the affected levels and modalities, several classic and other subtypes of aphasia have been defined, see Table 1. Nonetheless, these clear-cut classic types of aphasia do not frequently occur as the patient usually suffers from a combination of them (Bastiaanse, 2011).

The most well-known classic subtypes of aphasia are Broca's aphasia and Wernicke's aphasia. If a patient suffers from Broca's aphasia, then language production is impaired in the form of non-fluent speech

with a low speech rate. The characteristic of this type of aphasia is called telegram style, which entails that function words and grammatical morphemes are mostly left out, resulting in telegraphic speech. Language comprehension is relatively intact, whereas repetition of function words and sentences is impaired (Bastiaanse, 2011).

Wernicke's aphasia is characterized by phonemic and/or verbal paraphasias and/or neologisms, which is often called jargon. Paraphasias are instances where the patient makes mistakes at word-level. These errors can either concern the phonology of the word (i.e. phonemic paraphasias), such as *tegetable* instead of *vegetable* (Hallowell, 2017, p. 160) or the word itself when it is replaced by a semantically related, such as *ear* instead of *nose* (Hallowell, 2017, p. 160) or unrelated concept, such as *car* instead of *ladder* (Hallowell, 2017, p. 160) (i.e. verbal paraphasias). Sometimes the patient even utters words that are impossible to lead back to existing words (i.e. neologisms), such as *trunket* instead of *drink* (Bastiaanse, 2011; Hallowell, 2017, p. 160). Whereas the patient's fluency is relatively intact, his or her language comprehension is severely impaired, resulting in communicational difficulties. Besides impaired comprehension, the patient also suffers from impaired repetition (Bastiaanse, 2011). When compared to Broca's aphasia, patients that suffer from Wernicke's aphasia predominantly have problems at word-level with impaired comprehension whereas Broca's aphasics are mostly impaired at sentence level with impaired language production.

Other classic subtypes are anomic aphasia, conduction aphasia, transcortical aphasia (see next paragraph), and global aphasia. Language production and comprehension are relatively intact in anomic aphasia, but a patient suffering from this type of aphasia might encounter problems in the comprehension of longer and more complex sentences. The patient experiences wordfinding problems, which may result in pauses, circumlocutions, and empty speech. The patient tries to describe a word that he is unable to find without naming it (i.e. circumlocution) and the language use often lacks explicit information (i.e. empty speech) (Bastiaanse, 2011).

Conduction aphasia is characterized by phonemic paraphasias. The patient's comprehension and fluency are relatively unimpaired, however, due to self-correcting the phonemic paraphasias (i.e. conduite d'approche), the patient's speech rate and fluency might decrease (Bastiaanse, 2011).

The last and most severe subtype of aphasia is global aphasia where both language comprehension and production are impaired. A comorbidity of, amongst others, global aphasia is verbal apraxia. Verbal apraxia entails that articulation is impaired, which makes it impossible for the patient to plan the steps of the articulatory process (Bastiaanse, 2011). Although rare, verbal apraxia can also occur in the absence of aphasia (Bastiaanse, 2011).

A type of aphasia that in general is very rare, but that often occurs in brain tumor patients is dynamic aphasia. Dynamic aphasia is characterized by a spontaneous speech deficit and can be subdivided into pure dynamic aphasia and mixed dynamic aphasia (Robinson et al., 2006; Satoer et al., 2014). In case of pure dynamic aphasia, the patient's initiation of spontaneous speech is impaired with intact repetition, naming, and language comprehension, whereas a patient suffering from mixed dynamic aphasia can have additional impairments in one or more linguistic levels (Robinson et al., 2006; Satoer et al., 2014). Dynamic aphasia is further classified into language-specific and domain-general dynamic aphasia, the former involving impairments in word and sentence generation, whereas the latter extends to discourse (i.e. multiple responses) in both verbal and non-verbal communication (Robinson et al., 2006; Robinson et al., 2015). Suggested causes for dynamic aphasia are issues in conceptualization, involving concept selection and verbal planning (language-specific dynamic aphasia) and semantic strategy formation and/or use, the generation of verbal and/or non-verbal ideas, and fluent sequencing of novel thought (domain-general dynamic aphasia) (Robinson et al., 2015). In addition, executive and attentional

functions may be impaired, such as the ability to select important information to be communicated (i.e. task-setting), to maintain focus on the topic and the communicative intentions (i.e. energization), and to self-monitor utterances for suitability in the conversation (Robinson et al., 2015).

The term dynamic aphasia has been introduced to distinguish between two types of transcortical aphasia, namely transcortical sensory aphasia and transcortical motoric aphasia (i.e. frontal dynamic aphasia) (Nichelli, 2016; Robinson et al., 2015). A patient who suffers from sensory transcortical aphasia has a relatively intact fluency and repetition, whereas language comprehension is impaired and repetition can sometimes be restricted to echolalias and familiar texts. Echolalias are instances where the patient continuously repeats words or sentences that have been uttered by the interlocutor. The patient also speaks with neologisms and paraphasias, just as Wernicke's aphasics (Bastiaanse, 2011). A patient with motoric transcortical aphasia has intact comprehension and (a relatively) intact repetition, however, fluency is impaired. The patient often is not able to produce language himself, whereas he might be able to do so when the interlocutor supplies a sentence that the patient is expected to finish (Bastiaanse, 2011).

Table 1. Classic subtypes of aphasia and additional subtypes, based on Nouwens and colleagues (2013, p. 54), Grossman and Irwin (2018, p. 748), Robinson and colleagues (2006; 2015), and Satoer and colleagues (2014).

Aphasia type	Fluency	Comprehension	Repetition
Global aphasia	-	-	-
	Characteristics: all modalities and linguistic levels are severely impaired. Sparse language production that mainly consists of recurring utterances and automatisms and/or stereotypes.		
Broca's aphasia	-	+	-
	Characteristics: moderate to good comprehension, substantial language production issues with problems in the initiation of speech and phonemic inaccuracies. Many words are missing and/or not inflected (telegram style/agrammatism). Previously known as motor/expressive aphasia.		
Wernicke's aphasia	+	-	-
	Characteristics: language production consists of an abundance of inadequate function words and grammatical structures (paragrammatism) with semantic and phonemic paraphasias and neologisms. Pressured speech of which the message is incomprehensible. Impaired auditive input and thus impaired comprehension and unaware of mistakes. Previously known as sensory/receptive aphasia.		
Anomic aphasia	+[a]	+	+
	Characteristics: wordfinding problems. Mild anomic aphasia is characterized by word finding problems efficiently masked with synonyms or descriptions. Severe anomic aphasia is characterized by laborious, non-fluent, and slow speech with long thinking breaks and uh-interjections. It also frequently involves empty speech, such as *thing* and *that*.		
Conduction aphasia	+[a]	+	-
	Characteristics: fluent speech with conduite d'approche and phonemic paraphasias, mostly present in repetition.		
Transcortical sensory aphasia	+	-	+
	Characteristics: fluent speech, impaired comprehension, and repetition is relatively good (unlike Wernicke's aphasia).		

Transcortical motoric aphasia	-	+	+
	Characteristics: non-fluent speech, intact comprehension, and repetition is relatively good (unlike Broca's aphasia).		
Pure dynamic aphasia	-/+	**+**	**+**
	Characteristics: all linguistic levels are intact, except for spontaneous speech.		
Mixed dynamic aphasia	-/+	-/+	-/+
	Characteristics: spontaneous speech deficit with additional impairments in one or more linguistic levels.		

Note. '-' = impaired; '+' = relatively intact. -/+ = mildly impaired

[a] May be slowed by word finding difficulties and circumlocutions.

Symptoms and severity of aphasia can consist of impaired language production and/or perception of one or more language levels depending on the type, location and size of the lesion and in case of tumor-associated aphasia, the tumor grade, type, size, and location (Bastiaanse, 2011; Hallowell, 2017; Satoer et al., 2018). Patients with slow growing tumors often suffer from mild aphasia due to increased neural plasticity, where other brain regions can take over the functions of the damaged brain regions (Satoer et al., 2018; Satoer et al., 2019). Mild aphasia entails that the patient suffers from a low degree of aphasia with very subtle symptoms that go relatively unnoticed. These symptoms can manifest itself in word-finding, reading, writing, spontaneous speech and/or nuanced speech (Satoer et al., 2019). Apart from language deficiencies, people with mild aphasia often also experience additional difficulties with respect to cognitive abilities, emotions and behavior, their physical condition, and their position in society. Despite the term "mild", patients indeed suffer from the consequences of mild aphasia as it negatively influences their quality of life (Satoer et al., 2019).

To determine the aphasia type and severity, individuals with symptoms of aphasia as a consequence of a neurological disorder are often directed by a physician to an outpatient clinic of a hospital or rehabilitation center, or seek for help themselves in the chronic phase (≥ 1 year post onset). Symptoms of aphasia can be detected by a clinical linguist and/or a speech therapist using neurolinguistic tests. Based on their findings, a diagnosis can be made in order to inform the patient/proxy about the underlying disorder, to tailor a language therapy program and/or progress after speech therapy can be monitored. Early diagnosis and treatment is essential as it enlarges the chance of full recovery (Talacchi et al., 2013).

1.2 Diagnostic Instrument for Mild Aphasia (DIMA)

Usually classic standardized language tests, such as the Boston Naming Test (Kaplan et al., 1983) and the shortened Token Test (De Renzi & Faglioni, 1978) are used for diagnosing a patient with aphasia. However, these language tests alone are not sensitive enough for the diagnosis of patients with only a mild form of aphasia as most patients perform at ceiling level. This leads to underdiagnosis which leaves the patient with communicative difficulties in both daily life and at work (Satoer et al., 2019). A Dutch test that detects a mild form of aphasia did not yet exist. Therefore, Satoer and colleagues (2019) developed the Diagnostic Instrument for Mild Aphasia (DIMA).

The DIMA is based on the Dutch Linguistic Intraoperative Protocol (DuLIP) (De Witte et al., 2015), a very extensive linguistic test battery for brain tumor patients treated with awake surgery (Satoer et al., 2019).

The DIMA consists of four subtests, see Table 2: (1) repetition, which includes repetition items for 3-syllabic words (PHON A), compounds (PHON B), nonwords (PHON C), and sentences (PHON D); (2) odd picture out, which includes pictures of objects (SEM A) and actions (SEM B); (3) sentence completion (SYN A); and

(4) sentence judgement, which tests phonology (PHON E), semantics (SEM C), and syntax (SYN B). Some of these tasks are revised DuLIP tasks.

Table 2. Task description and examples of the DIMA (Satoer et al., 2019, p. 5)

Linguistic level	Task	Stimuli (examples)
Phonology	Repetition:	
	· PHON A: 3-syllabic words	
	alternating stress patterns	*gorilla* ('gorilla'), *oefening* ('practice')
	phonemic similarities	*domino* ('domino')
	consonant clusters	*constructie* ('construction')
	· PHON B: syllabic compounds	
	alternating stress patterns with phonemic similarities and/or consonant clusters	*opdrachtgever* ('client'), *ontdekkingsreiziger* ('explorer'), *afstandbediening* ('remote control') *schouderblad* ('shoulder blade')
	· PHON C: syllabic nonwords	
	alternating stress patterns with phonemic similarities and/or consonant clusters	inima, anato kikira trendelorig
	· PHON D: sentences	
	alternating length with phonemic similarities and/or consonant clusters	*Harry rent zich rot* – 4 words ('Harry is running fast'), *Iedere vrijdag eten we vreselijke vissticks* – 6 words ('Every Friday we eat terrible fishsticks'). *De Griek ontdekte vier nietjes in de band van zijn fiets* – 11 words ('The Greek discovered four staples in the tire of his bike'). *In welke la leg jij de lepel?* ('In which drawer do you put the spoon?'). *De prikkelbare panter besprong haar prooi* ('The irritable panther jumped its prey').
	PHON E: Sentence judgment:	
	· phonologically (in)correct sentences	*De wikker gaat naar parot* ('The wikker goes to parot').
Semantics	Semantic odd picture out:	
	· SEM A: objects	*slang, hond, kat* ('snake, dog, cat')
	· SEM B: actions	*scheren, douchen, baden* ('shave, shower, bathe')
	SEM C: Sentence judgment:	
	· semantically (in)correct sentences	*De sigaar verveelde zich* ('The cigar was bored').
Morphosyntax/ grammar	SYN A: Sentence completion	*Hij viel van…* ('He fell from…'), *Om 5 uur…* ('At 5 o'clock…')
	SYN B: Sentence judgment	
	· Syntactically (in)correct sentences	*Hij gaat geschilderd op de muur* ('He goes painted on the wall').

The phonological repetition task assesses "word production according to the phonological input- and output route" (Satoer et al., 2019, p. 5). The patient is instructed to repeat a word or a sentence that is uttered by the clinical linguist. This can either be a three-syllabic word, a compound word, a nonword, or a sentence. To increase the level of difficulty, which facilitates the detection of both mild and severe

aphasia, the words and sentences each contain alternating stress patterns, phonemic similarities, and/or consonant clusters. The patient's answer is scored 1 in case of a correct answer that has been given within four seconds. The answer is considered incorrect (0) when the patient gives the wrong answer, is hesitant, self-corrects, or does not answer within the timeframe of four seconds (Satoer et al., 2019). The maximum score for this task is 40 as ten points can be earned in each of the subtests.

During the sentence judgment task that assesses phonological, semantic, and grammatical awareness, the test leader visually presents the patient randomized correct and incorrect sentences on a laptop screen. The patient is asked to judge whether these sentences are correct at the levels of phonology (PHON E), semantics (SEM C), and morpho-syntax (SYN B). A correct answer is scored 1, whereas an incorrect answer is scored 0. Additionally, the patient's reaction time in milliseconds is recorded.

For the semantic odd picture out objects and actions that assesses "naming and non-verbal semantic judgment" (Satoer et al., 2019, p. 5), the patient will be shown a series of three black and white pictures that either denote objects, animals, or actions. All pictures are semantically related, except for one of them. The patient is asked to name the picture that does not fit within the series. The patient's response is scored correct (1) if the correct answer is given within the timeframe of four seconds and incorrect (0) when he or she exceeds the timeframe, hesitates, self-corrects, or gives the incorrect answer (Satoer et al., 2019). This subtest has a maximum score of ten as there are five series of pictures for both verbs and actions.

To conclude, the sentence completion task assesses spontaneous speech in context provided by the clinical linguist. The patient is asked to finish the sentence in a way that is syntactically and semantically appropriate with either a word or a constituent (Satoer et al., 2019). Some of the sentences allow more diversity in the answers than others. Answers are scored correct (1) if the sentence is finished fluently, in a meaningful way, and is initiated within the timeframe of four seconds. Responses that do not adhere to these criteria are scored 0 and there is a maximum score of ten as the subtask consists of ten items.

The DIMA was validated in a population of 214 healthy individuals, see Appendix A for the demographic characteristics, and is a potentially valid diagnostic tool for the detection of mild aphasia (Satoer et al., 2019). The semantic odd picture out actions task (SEM B) and the phonological tasks (PHON A-E) still require some improvement as the healthy control group did not perform at ceiling level (SEM B) and the Dutch norm group outperformed the Flemish norm group (PHON A-E) (Satoer et al., 2019). Furthermore, a questionnaire and other tests, such as reading and writing, might be included later to have a subjective measure and to cover more symptoms of mild aphasia (Satoer et al., 2019). Nevertheless, subtests of the DIMA have been a starting point for the development of the Dutch version of the TeleLanguage Test (De Witte et al., 2019).

1.3 TeleLanguage Test

The TeleLanguage Test is a telephone-based language test that is comprised of two versions, A and B. Both versions consists of the following subtests that are conducted in this order, see Table 3.

Table 3. Subtests of the TeleLanguage Test (De Witte et al., 2019, p. 95)

Task	Example
Comprehension screening (adapted from WAB)	*I am going to ask you some questions. Answer yes or no:* *1. Is your name X?* *2. Do you eat a banana before you peel it?*

Verbal naming test (adapted from verbal naming test[1])	*I am going to describe an object or a verb and I want you to tell me the name of what I am describing:* *1. A large animal in Africa with a trunk.* *2. What ice does when it gets hot.*
Word and sentence repetition (adapted from WAB, BDAE)	*Repeat after me:* *1. Bed* *2. Screwdriver* *3. Methodist episcopal* *4. He unlocks the heavy oak door.*
Letter fluency (adapted from WAB)	*Name as many words that start with the letter F as you can in 1 minute.*
Semantic noun and verb selection (adapted from DuLIP)	*Two words will go together best because of their meaning and one word will not. Tell me the word that does not fit.* *1. Banana, apple, carrot* *2. Talk, tell, sing*
Semantic fluency (adapted from WAB)	*Name as many animals as you can in 1 minute.*
Story completion (adapted from Goodglass story completion[2])	*Complete the story:* *1. My dog is hungry and I have a bone in my hand. What's next?* *2. The mouse was running around. A cat came along. The mouse did not see the cat running after him. What happened to the mouse? The mouse…*
Sentence completion (adapted from DuLIP)	*Complete the sentence in a meaningful way:* *1. The man knows that…* *2. I'll do it when…*

The comprehension screening consists of five questions that can be answered by either *yes* or *no* to assess the participant's comprehension. This task has been included to exclude participants with a comprehension deficit (i.e. less than four correct answers) as this makes it impossible to conduct a the telephone language assessment.

The comprehension screening is followed by the verbal naming test. This test consists of 25 descriptions of or questions leading to a target word that the participant is asked to name.

The third test in the protocol is the word and sentence repetition test that assesses verbal repetition. The test consists of a total of ten words and sentences which are given by the test leader and should be repeated by the participant.

The repetition test is followed by letter fluency, which assesses phonemic fluency. The patient is requested to name as many items from the same phonemic category (i.e. starting with the same letter) within the timeframe of one minute.

The fifth test in the protocol is semantic noun and verb selection, which assesses verbal semantic judgment, semantic processing, naming, and semantic knowledge (De Witte et al., 2015). The test leader lists three nouns or verbs of which one does not semantically fit. The patient is asked to name the noun or verb that does not fit.

The semantic noun and verb selection task is followed by semantic fluency. This task requires the patient to name as many animals from the same semantic category (i.e. all of the category 'animal') within one minute.

[1] Yochim et al. (2015).
[2] Goodglass & Kaplan (1972).

The seventh test in the protocol is story completion, which assesses spontaneous speech. The test consists of five items that contain beginnings of a story. The test leader provides these story beginnings to the patient and then asks him to complete the story.

The test ends with sentence completion, which assesses spontaneous speech in context. The test leader provides the participant with the beginning of a sentence which the patient is asked to finish in a meaningful way. The test consists of a total of ten items.

All tests have practice items and an increasing level of difficulty in order to identify the degree of language impairment. All items have been selected based on word frequency, word length, and morphological/phonological form. A single repetition is allowed if requested and minor dysarthric errors are allowed. Administration of the test should take no longer than 20 minutes.

The TeleLanguage Test appeared to be a feasible and valid test battery as an alternative for face-to-face assessment (De Witte et al., 2019). Both versions of the test proved to be reliable for detection of mild aphasia. Healthy participants performed at ceiling level during most tests except for the fluency tasks and some tasks (e.g. verbal naming and semantic selection) turned out to be more sensitive when administered by phone than in-person. Limitations are the lack of visual input, which makes administration and analysis of tasks such as word and sentence repetition more challenging. For this reason, nonwords – although sensitive – have been excluded from the test battery. Furthermore, external factors such as environmental noise and internal factors such as tiredness cannot be fully controlled (De Witte et al., 2019).

1.4 TeleTaalTest-NL (TTT-NL)

Where the TTT-NL was originally designed for postoperative assessment of language abilities in glioma patients who underwent awake craniotomy, the test might also be applicable to patients with a neurological language disorder of other origins, such as stroke or neurodegenerative diseases (e.g. Alzheimer's disease) (Satoer et al., 2020a). The TTT-NL will never be able to replace face-to-face assessment since the observations made are essential for the determination of the right diagnosis, but in some situations, such as COVID-19, no possibility for hospital visits due to long travel distances, or no access to video calling, telephone assessment is the only option. This way of testing offers a possibility to screen and monitor patients from a distance to ensure that patients and their partners can still get advice and possibly therapy from a distance (Satoer et al., 2020a).

The goal of the TTT-NL is to indicate mild to severe language disorders by means of telephone assessment. The TTT-NL originally consisted of a comprehension task and six additional subtests (now brought down to four subtests, see Table 4), which are the same tasks as those included in the TeleLanguage Test (Ruijs, 2018; Satoer et al., 2020a; Van Dijk, 2018). Each of these tasks evaluates different linguistic levels and modalities, which can be impaired in different symptoms of aphasia (Van Dijk, 2018). The comprehension task consists of five questions and is there to filter out patients with impaired comprehension as intact comprehension is necessary to successfully complete the DTLB. In case comprehension is intact, the test leader will administer a word retrieval test (20 items; fifteen nouns, five verbs), phonological repetition (ten items; six words, four sentences), phonological fluency (one letter), semantic odd word out (ten items), semantic fluency (one category), and sentence completion (ten items). All subtests except for word retrieval, which has been developed by the authors themselves, have been based on the DuLIP or DIMA (Ruijs, 2018). Two versions have been developed to prevent a learning effect in postoperative assessments. The subtests and their order are the same, but the items are different.

Table 4. The development of the TeleTaalTest-NL.

Subtest	**DTLB** (Ruijs, 2018) version A & B	**TeleLanguage NL** (Van Dijk, 2018) version A & B	**TTT-NL** (Satoer et al., 2020a) one version (version A Van Dijk, 2018)
Comprehension	5 items	5 items version A: *4. Komt de maand maart voor juni?* 'Does the month March precede June?' → *Komt de maand mei voor juni?* 'Does the month May precede June?' version B: *Bent u een dokter?* 'Are you a doctor?' → *Bent u een tandarts?* 'Are you a dentist?'	5 items *5. Is een vlinder een vogel* 'Is a butterfly a bird?' → *Is bloemkool een fruitsoort?* 'Is cauliflower a type of fruit?'
Word retrieval (adapted from TeleLanguage Test)	20 items (15 nouns, 5 verbs)	20 items (17 nouns, 3 verbs) version A: *school* 'school' → *school/lokaal* 'school/classroom' *wonen* 'to live' → *dansen* 'to dance' *vissen* 'fish' → *films* 'movies' *kasteel* 'castle' → *kasteel/paleis* 'castle/palace' *scheren* 'to shave' → *hamer* 'hammer' *Wie bakt er brood?* 'Who bakes bread?' → *Wie maakt er brood?* 'Who makes bread?' *kledingstuk* 'garment' → *woonboot* 'houseboat' *duinen* 'dunes' → *eskimo* 'Eskimo' version B: *politie* 'police' → *zoon* 'son' *Amsterdam is een …?* 'Amsterdam is a …?' → *Zwolle is een …?* 'Zwolle is a …?' *dokter* 'doctor' → *kapitein* 'captain' *lezen* 'to read' → *verlaten* 'to abandon' *Wat doet een vliegtuig?* 'What does an airplane do?' → *Zich met vleugels door de lucht voortbewegen noem je ook wel…?* 'To move forward through the air with wings is also called …?' *brug* 'bridge' → *keuken* 'kitchen' *strand* 'beach' → *restaurant* 'restaurant' *ochtend* 'morning' → *papier* 'paper' *sigaret* 'cigaret' → *siga(a)r(et)* 'cigar/cigaret' *koffers* 'suitcases' → *angel* 'sting'	20 items (17 nouns, 3 verbs) *1. vader* 'father' → *oma* 'grandmother' *3. Waar zitten je oren aan vast?* 'To what are your ears attached?' → *Je oren zitten aan je …?* 'Your ears are attached to your …?' *4. school/lokaal* 'school/classroom' → *school* 'school' *12. Welk drankje komt er van de koe?* 'Which drink comes from the cow?' → *Welke drank komt er van de koe?* 'Which beverage comes from the cow?' *14. kasteel/paleis* 'castle/palace' → *kasteel* 'castle' *16. Waarmee sla je een spijker?* 'With what do you hit a nail?' → *Waarmee sla je een spijker in de muur?* 'With what do you hit a nail into the wall?' *17. Hoe noem je een baby van een hond?* 'What do you call the baby of a dog?' → *Hoe noem je een jong van een hond?* 'What do you call the young of a dog?' *18. Wie maakt er brood?* 'Who makes bread?' → *Waar worden broden en taarten gemaakt?* 'Where are bread and cakes being made?' *20. Hoe noem je mensen die op de noordpool wonen?* 'What do you call people who live at the north pole?' → *Hoe noem je mensen die dicht bij de noordpool wonen?* 'What do you call people who live close to the north pole?'

		zinken 'to sink' → *diploma* 'diploma' *kieuwen* 'gills' → *lifter* 'hitchhiker' *kitten* 'pussycat' → *pubertijd* 'puberty'	
Phonological repetition (adapted from DIMA)	10 items (6 words, 4 sentences)	10 items (6 words, 4 sentences) version B: *slanugári → froduslóka*	10 items (6 words, 4 sentences) *De jongen vergat zijn vissen te voeren.* 'The boy forgot to feed his fish.' → *De vrouw vergat de vissen te voeren.* 'The woman forgot to feed the fish.'
Phonological fluency	1 letter, 1 minute	1 letter, 1 minute	-
Semantic odd word out/semantic selection (adapted from DuLIP)	10 items	10 items version A: *bal pop zeep* 'ball doll soap' → *kaas, boter, pet* 'cheese, butter, cap'	10 items
Semantic fluency (adapted from DuLIP)	1 category, 1 minute	1 category, 1 minute	-
Sentence completion (adapted from DIMA)	10 items	10 items	10 items

A first validation of the DTLB was carried out by Ruijs (2018) on 30 healthy controls and six former patients who have undergone awake craniotomy. It revealed that the total scores and the scores of most subtests of versions A and B did not differ from each other significantly, except for phonological repetition, where the repetition of nonwords proved to be difficult. This finding is supported by Sierpowska and colleagues (2017) who indicate that, although more sensitive than existing words in a repetition task, nonwords appear to be more difficult and prone to incorrect repetitions, even in healthy adults. Additionally, some items of the word retrieval task had multiple correct answers. Despite these suggested revisions, Ruijs (2018) concluded that the DTLB is a valid tele-language test that should be further developed.

Later on, the DTLB was improved and developed into the TeleLanguage-NL (Van Dijk, 2018). Some revisions were made in both versions of the comprehension screening, phonological repetition and the semantic odd word out, whereas many items of the word retrieval subtest were revised, resulting in a division of seventeen nouns and three verbs. After administering the TeleLanguage-NL in 28 healthy individuals, Van Dijk (2018) concluded that the test is valid for multiple clinical and scientific purposes. Further development of word finding is advised as some items had multiple correct answers or descriptions needed improvement, and the number of nouns and verbs is unbalanced. Furthermore, the test had not been compared to in-person assessment of a similar test as in the study by De Witte and colleagues (2019).

The most recent version of the TTT-NL has been developed by Satoer and colleagues (2020a). This version consists of only one version of the test that has been developed by Van Dijk (2018). In the comprehension screening, item 5 *Is een vlinder een vogel?* 'Is a butterfly a bird?' has been replaced by *Is bloemkool een fruitsoort?* 'Is cauliflower a type of fruit?' as this is an item from the Afasie Bedside-Check (ABC) (Visch-Brink & El Hachioui, 2013). Furthermore, several items in the word finding task have been revised. The target word for item 1 has been changed from *vader* 'father' to *oma* 'grandmother' as the previous questions was outdated. The other changes made in the word finding task all concern descriptions (i.e. items 3, 12, 16-18, and 20) and accepted answers (i.e. items 4 and 14). Finally, one item in the phonological repetition task was changed from *De jongen vergat zijn vissen te voeren* 'The boy forgot to feed his fish' to *De vrouw vergat de vissen te voeren* 'The woman forgot to feed the fish' as item 7 already contained *jongens* 'boys' and the new item includes alliteration, which makes the item more challenging. Provisional norm data for this test of a control group of 28 healthy participants (11 female, 17 male), aged 40-70 with differing levels of education (Verhage 2-7), have been established by Satoer and colleagues (2020a) and can be found in Table 5. The cutoff scores with regard to the norm data are a standard deviation (SD) of ≥1.5 for a score that is below average and an SD of ≥2 for a score that is deviant.

Table 5. Provisional norm data for the TeleTaalTest-NL (Satoer et al., 2020a, p. 5)

Subtest (total score)	Mean	SD	Range
Word finding (40)	38.82	1.25	35 - 40
Phonological repetition (18)	17.96	0.19	17 - 18
Semantic odd word out (20)	19.96	0.19	19 - 20
Syntax (20)	19.75	0.59	18 - 20

The development of the TTT-NL is still ongoing. The present study contributes to the development and validation of the TTT-NL by improving its items and subtests and supplementing its norm data with at

least 20 healthy individuals for a more reliable reference group. External validity is assessed by comparing TTT-NL scores to previously collected data of the DIMA.

In the present study, the following research questions are addressed:

1. Is the TeleTaalTest-NL (Satoer et al., 2020a) a valid tele-language test when administered in a healthy control group?
2. Do the TeleTaalTest-NL (Satoer et al., 2020a) and its subtests correlate with the DIMA (Satoer et al., 2019) and its subtests?

To answers these questions, the current version of the TTT-NL is first revised and assessed in a pilot control group of ten healthy individuals in order to select the most optimal items for the subtests based on an 80% rate of correctness. Based on these results, a new version of the TTT-NL is constructed. Next, the new version of the TTT-NL is administered in a healthy control group of at least 20 participants, that serves as a baseline for language performance. To assess the construct validity of the TTT-NL, the healthy control data from the TTT-NL is then compared to the healthy control data from the DIMA.

Both healthy control groups are described in section 2.1, the test materials, amongst others the TTT-NL, of which both the pilot and final versions are outlined in section 2.2.2, can be found in section 2.2. The tests are followed by the full design and procedure (2.2.3), the outcome measures (2.2.4), and the statistical analysis of both the pilot and the final version (2.2.5). Results on the pilot study and final version of the TTT-NL are reported in section 3. In conclusion, the results are discussed (4.1), its clinical implications (4.2) and limitations reported (4.3), and concluding remarks can be found in section 5.

Chapter 2: Methodology

2.1 Participants

For the selection of items for TTT-NL (Satoer et al., 2020a) we first recruited a pilot control group of twelve native Dutch healthy controls (BA and MA students) of which the age, gender, educational level on both Verhage-scale (Verhage, 1964) and education in years, and handedness assessed with the Edinburgh Handedness Inventory (EHI) (Oldfield, 1971) were collected in conjunction with the administration of the TTT-NL.

The study group for the administration of the final version of the TTT-NL consists of 31 healthy controls of which the age, gender, educational level on both Verhage-scale (Verhage, 1964) and education in years, and handedness assessed with the EHI (Oldfield, 1971) were collected in conjunction with the administration of the TTT-NL. The current study strived for an equal number of both genders and a diversity of ages and educational levels in order to provide a representative norm group. In order for the data to be comparable to the previously gathered norm data on the DIMA-study (Satoer et al., 2019), participants have been split into two age groups in accordance with the presumed age of onset of cognitive decline, starting at age 55 (Rönnlund et al., 2005 in Satoer et al., 2019). This leads to one age group of age 18 to 54 and another age group of age 55 to 74. The same applies to the years of education as the minimum, compulsory period of education in the Netherlands (i.e. primary school and high school until age 18) equals twelve years (Satoer et al., 2019).

In case of both healthy control groups, participants must be native speakers of Dutch and 18 years of age or above. Exclusion criteria are: Neurological disorders affecting language (e.g. Alzheimer's disease), other language-related disorders (e.g. developmental disorders such as dyslexia and Autism Spectrum Disorder), uncorrected hearing deficiencies, phone call of which poor connection caused communication problems, and participants who have not been exposed to Dutch during most of their childhood. All participants have received an information letter (Appendix B) and gave their informed consent (Appendix C) prior to telephone assessment. This study was approved by the Faculty Ethics assessment Committee of the Faculty of Humanities (FEtC-H) at Utrecht University.

2.2 Materials

2.2.1 Edinburgh Handedness Inventory (EHI)

The Dutch adaptation of the Edinburgh Handedness Inventory (EHI) (Oldfield, 1971) was included to quickly and reliably assess the participants' handedness by means of ten questions, see Appendix D. The test leader asked the participant to indicate his preferred handedness in ten activities that involve holding objects, such as *schrijven* 'to write' and *tandenpoetsen* 'to brush your teeth'. In case more than one hand is involved in the activity, the hand in question is indicated between brackets (i.e. *lucifer aansteken (lucifer)* 'to light a match (match)'). The degree of preference is indicated by the number of plusses in the column for left and/or right. A strong preference (i.e. the participant would never use the other hand unless forced to do so) being indicated by two plusses in either the left or the right column, a mild/normal preference by one plus in one of the two columns, and no preference is indicated by one plus in both columns. The number of plusses are added up per column and they are put in the formula below the table to calculate a percentage for handedness. Scores below -40% are interpreted lefthanded, scores between and including -40% and 40% are interpreted ambidextrous, and scores

above 40% are interpreted righthanded. Administration takes five minutes at most and can be done by telephone.

In the pilot group, participants filled out the form themselves, whereas the EHI in the final test group was administered by phone before administration of the TTT-NL (Satoer et al., 2020a).

2.2.2 TeleTaalTest-NL

2.2.3 Design and procedure

The procedure is the same for the pilot study with student participants on the suitability of the proposed items for the TTT-NL (Satoer et al., 2020a) subtests that preceded administration of the test in the study group of 31 participants.

Data was collected by means of a telephone assessment. Prior to assessment, participants were approached via email. This email included an information letter (Appendix B) and an informed consent form (Appendix C), which needed to be signed and returned. Then, the date and time for the telephone assessment of the TTT-NL were selected.

The participant was requested to sit in a quiet room without any distractions. Preceding test administration, the participant was asked questions about his age, gender, education, profession, and hearing. Before actual assessment, the participant first received a general explanation about the test and its procedure.

Test administration was audio recorded using Audacity® (Audacity Team, 2020) to facilitate scoring. The TTT-NL consists of a pretest, and four subtests. Each subtest was preceded by a short instruction. Scoring was indicated on the form, which can be found in Appendices E & F. The TTT-NL consists of a pretest, which is a comprehension screening, and four subtests. All subtests contain an introductory question to practice what is expected from the participants.

Pretest: Comprehension screening - five items

The pretest assesses comprehension and consists of five short questions that can be answered by either *yes* or *no*. Each correct answer is scored 1, whereas each incorrect answer or no answer is scored 0. A maximum score of five can be obtained. This first test is essential, since all questions must be answered correctly (i.e. comprehension must be intact) to proceed to the other subtests.

Word finding - ten items / ten spare items

The first subtest is a word finding test. The test-leader will ask ten, or if necessary twenty questions (target words per ten items: five verbs, five nouns) that can, and should as often as possible, be answered with one word. Depending on the correctness and the time passed, the answers are scored from 0 to 2. A wrong answer is scored 0, a correct answer initiated after four seconds is scored 1, and a correct answer initiated within the timeframe of four seconds is scored 2. If the participant fails to give correct answers five times in a row, this subtest is aborted. The second set of ten items has been included as a backup that can be assessed in addition to the first set in case there is doubt about the results of the first ten items.

Phonological repetition - ten items

The phonological test assesses verbal repetition and is divided into three-syllabic words (two items), compounds (two-items), non-words (two items), and sentences (four items) with an increasing level of difficulty. The participant's answers are scored depending on the number of content words, which means that it can be 0, 1, 2, 3, or 4. Possible dysarthric and speech-apraxic answers cannot be distinguished without seeing the participant's articulatory movements (Satoer et al., 2020a).

Semantic selection - ten items

The third test is a semantic selection test that assesses semantic judgment. The experimenter presents the participant ten sequences of three words of which one does not semantically fit within the sequence. Depending on the participant's answer and explanation for the given answer, the items are scored 0, 1, or 2. A wrong answer is scored 0, a correct answer without explanation is scored 1, and a correct answer with explanation is scored 2.

Sentence completion - ten items

The final subtest is a verbal sentence completion task, where the experimenter gives the participant the beginnings of ten sentences that have to be finished by the participant. Depending on the speech fluency and the syntactic and semantic correctness of the sentence, the answers will be scored 0, 1, or 2. A syntactically incorrect sentence is scored 0, a sentence that contains hesitation, doubt, or repetition is scored 1, and a fluent and syntactically correct sentence is scored 2..

This leads to a maximum score of five for the comprehension screening, either 20 or 40 for the word finding task, 18 for the verbal repetition task, 20 for the semantic selection task, and 20 for the sentence completion task. Therefore, the minimum score on the TTT-NL is 0 and the maximum score amounts to 83 in case of ten items in the word finding task and 103 in case of two sets of ten items in the word finding task.

After tele-testing, the experimenter provided a rating about the quality of the phone call and wrote down additional comments about the test administration. Test administration took approximately 15 to 20 minutes.

2.2.3.1 Pilot version – item selection

The aim of the pilot study was to evaluate the old and new items for their rate of correctness, which had to be at least 80%, and to select items for the final version of the TTT-NL (Satoer et al., 2020a). All pre-existing items were evaluated for their timelessness (i.e. Will this item still be valid in for example 10 years from now or is it outdated?), effectiveness (i.e. Does it test what we would like to test, and for the word finding test, does the sentence lead to the target word?), and structure (i.e. Is the balance between items in terms of word class equal? Is the number of items logical and sufficient?). The pilot version of the test contained more items per subtest than intended for the final version to make exclusion of unsuitable items possible, see Appendix G. The full pilot version of the TTT-NL can be found in Appendix E. Results of the pilot study can be found in section 3.1.

2.2.3.2 Test version – norm study

Apart from the semantic selection task, which remained unchanged, and the word finding task that was revised to a greater extent, the comprehension screening and subtests underwent minor revisions that can be found in Appendix H. Furthermore, additional remarks on the number of repetitions by the experimenter (i.e. repetitions can be marked in a column) and the answer given by the participant (i.e. the last answer counts) have been added to all scoring sections of the screening and subtests. The complete final version of the TTT-NL can be found in Appendix F.

2.2.4 Outcome measures

The primary outcome measures for telephone assessment are the sub-scores and total score from the TTT-NL (Satoer et al., 2020a), concerning several language abilities. The pretest measures auditive language comprehension. The wordfinding task measures wordfinding, the phonological repetition subtest measures language production on the basis of the phonological in- and output route, the semantic selection test measures verbal semantic knowledge, and the sentence completion task measures spontaneous speech in context production (a combination of phonology, semantics, and syntax).

2.2.5 Analysis

To see if the proposed changes yield results with an 80% rate of correctness and to see if further improvement of the items is necessary, a descriptive analysis of the individual data from the pilot control group was performed. To examine the relationship between demographic variables and TTT-NL scores, Kruskal-Wallis tests and Mann-Whitney U tests were performed. To answer the question if the TTT-NL (Satoer et al., 2020a) is a valid tele-language test that measures similar language abilities when administered in a healthy control group, the healthy control data collected from the tele-test was compared to the healthy control data collected from the DIMA (Satoer et al., 2019) by means of Spearman correlation tests. Besides performing a correlation test on the total scores, results from the phonological repetition task of the TTT-NL have been compared to the total scores from the phonological repetition task (PHON A, B, C, D) of the DIMA. Furthermore, the results from the semantic selection task of the TTT-NL have been compared to the results from the semantic picture out nouns (SEM A) of the DIMA and the results from the sentence completion task of the TTT-NL have been compared to the sentence completion task (SYN A) of the DIMA. As the DIMA does not include a word finding task, the results from the word finding test have been compared to *z*-scores based on age-matched BNT norm data (Kaplan et al., 2001). As successful completion of the comprehension screening is only a criterium for further testing, no further correlations have been carried out. All analyses were performed using IBM SPSS Statistics 26 (IBM Corp., 2020).

Chapter 3: Results

3.1 Pilot version – item selection

After exclusion in concordance with the previously provided criteria (Autism Spectrum Disorder, n = 2), ten participants were included in the pilot study of the TTT-NL. The demographic characteristics of the healthy control group have been summarized in Table 6.

Table 6. Demographic characteristics of the healthy control group for the pilot study.

Demographics	Mean	SD	Range
Age in years	23.52	1.59	20.71 - 26.86
Education in years	17.40	1.58	15.00 - 21.00
Education Verhage	7	0	7
Demographics	Groups	Number of participants	Percentage
Gender	Female	8	80
	Male	2	20
Education level	≤ 12 years	0	0
	> 12 years	10	100
Education Verhage	7	10	100
Handedness	Left	1	10
	Right	9	90
	Ambidexter	0	0
Hearing	Good	10	100
	Moderate	0	0
	Poor	0	0
Hearing device	Yes	0	0
	No	10	100
Cell phone handedness	Left	1	10
	Right	3	30
	Missing	6	60

After administration of the pilot version of the TTT-NL, the items were checked for their rate of correctness and alternative answers given in order to see if items would have to be changed or excluded before administration of the final version of the TTT-NL.

Pretest: Comprehension screening

Both alternatives for item 2 had a 100% rate of correctness, just as all other items included in this test. Eventually, opt2 *Bent u een kind?* 'Are you a child?' was included in the final version of the TTT-NL since this question matches the length of the original question.

Word finding

Besides the items with less than an 80% rate of correctness in Table 7, items 19 and R1 have been excluded due to a priming effect (*schaatsers* 'ice skaters' for *schaatsbaan* 'ice rink') or because of too many possible alternatives for the target word (*munten* 'coins', *papier geld* 'paper money/cash' are all types of money). The questions for items R.6 (Table 8) and R.7 have been modified after having tested

eight participants. The reason for these adjustments were the wrong assignment of thematic roles (*voeden* 'to feed' implies that someone is feeding the ducks rather than that the ducks are being fed) and the incorrect interpretation of the question (*opletten* 'to pay attention' suggests a classroom setting whereas the target word refers to alerting someone in a possibly threatening situation). In case of R.7V, the adjustment did not yield better results, which is why the item will not be included in the final test.

Apart from the excluded items in Table 7, multiple answers were given to the items in Table 8. Nevertheless, these items still have an 80% or above rate of correctness and they thus remained included for the final round of testing.

Table 7. Excluded word finding test pilot items.

Word finding				
Item	Question	Target word	Alternative answers given	Rate of correctness
III.3	*Hoe noem je dat als je iets even niet meer weet?* 'What do you call it when you do not remember something for a moment?'	*vergeten* 'to forget'	*nadenken/denken* 'to think' (3); *vergeten zijn* 'to have forgotten' (1, correct); *black-out* 'blackout' (1)	60%
R.7[3]	*Hoe noem je het als je zegt dat iemand moet opletten?* 'What do you call it when you say that someone has to watch out?'	*waarschuwen* 'to warn'	*bij de les halen/houden* 'to pay attention to the lesson/to be alert' (2); *aandacht erbij houden* 'to pay attention' (1); *weet ik niet* 'I do not know' (1)	66.6%
R.7V[4]	*Hoe noem je het als je tegen iemand zegt dat hij moet opletten?* 'What do you call it when you tell someone to watch out?'	*waarschuwen* 'to warn'	*op de vingers tikken* 'to rap someone's knuckles' (1); *terecht wijzen* (1); *weet ik niet* 'I do not know' (2)	0%
19	*Hoe noem je de plek waar schaatsers trainen?* 'What do you call the place where ice skaters train?'	*ijsbaan* 'ice rink'	*schaatsbaan* 'ice rink' (2, correct)	100%
R.1	*Contant betalen doe je met …?* 'You pay in cash with …?'	*geld* 'money'	*munten* 'coins' (1); *papier geld* 'paper money/cash' (1)	80%

Table 8. Included word finding test pilot items to which alternative answers were given.

Word finding				
Item	Question	Target word	Alternative answers given	Rate of correctness
8	*Bij een wedstrijd wint het ene team, wat doet het andere team?*	*verliezen* 'to lose'	*verliest* 'loses' (2, correct)	100%

[3] Administered in eight participants.
[4] Adjustment of R.7, administered in four participants.

	'At a game, one team wins; what does the other team do?'			
11	*Wat draaien ze in de bioscoop?* 'What do they show in the cinema?'	*films* 'movies'	*film* 'movie' (5, correct)	100%
14	*Waarin wonen een prins en prinses?* 'Where do a prince and princess live?'	*kasteel* 'castle'	*paleis* 'palace' (4, correct); *kasteel of paleis* 'castle or palace' (1, correct)	100%
18	*Waar worden broden en taarten gemaakt?* 'Where are bread and cakes being made?'	*bakkerij* 'bakery'	*bakker* 'baker' (4, correct)	100%
II.14	*Wat ben je doen als je getallen op volgorde opnoemt?* 'Wat are you doing when you are naming numbers in order?'	*tellen* 'to count'	*opsommen* 'to enumerate' (1); *weet ik niet* 'I do not know' (1)	90%
II.19	*Wat doet iemand die mensen in trance brengt?* 'What does someone who entrances people do?'	*hypnotiseren* 'to hypnotise'	*illumineren* 'to illuminate' (1)	90%
II.20	*Hoe noem je het wanneer een tuinman takken van een boom knipt?* 'What do you call it when a gardener cuts branches off a tree?'	*snoeien* 'to prune'	*tuinwerken* 'to work in the garden' (1)	90%
R.6[5]	*Hoe noem je het wanneer eenden stukjes brood krijgen?* 'What do you call it when ducks are being fed pieces of bread?'	*voeren/voederen* 'to feed' (BE)	*weet ik niet* 'I do not know' (1)	83.3%

Phonological repetition

All words, non-words and sentences had a 100% rate of correctness, apart from *frimótika*, see Table 9.

Table 9. Pilot items in the phonological repetition test with less than an 80% rate of correctness.

Phonological repetition			
Item	Target word	Alternative answers given	Rate of correctness
6	*frimótika*	*imozika* (1), *imotika* (2), *vri/fri/primotika* (1), *timotika* (1)	50%

[5] Administered in the first eight participants.

Semantic selection

All items in this subtest had a 100% rate of correctness. Hence, no further changes have been made for the final version of the TTT-NL.

Sentence completion

All items, except for item 1, had a 100% rate of correctness. No further changes have been made for the final version of the TTT-NL, except for a clarification of the description of *twijfel* 'doubt'.

See Appendix H for a more elaborate description of the item selection for the test version of the TTT-NL.

3.2 Test version – norm study

After exclusion in concordance with the previously provided criteria (epilepsy, n = 1; poor connection, n = 1; underexposed to Dutch during childhood, n = 1), 28 participants were included in the final administration of the TTT-NL. The demographic characteristics of the healthy control group are summarized in Table 10.

Table 10. Demographic characteristics of the healthy control group for the final study.

Demographics	Mean	SD	Range
Age in years	43.17	19.14	18.56 - 74.98
Education in years	16.36	4.06	10 - 28
Education Verhage	6.25	0.89	4 - 7
Demographics	**Groups**	**Number of participants**	**Percentage**
Gender	Female	13	46.4
	Male	15	53.6
Age group	18-54 years	17	60.7
	55-74 years	11	39.3
Education level	≤ 12 years	4	14.3
	> 12 years	24	85.7
Education Verhage	4	2	7.1
	5	2	7.1
	6	11	39.3
	7	13	46.4
Handedness	Left	4	14.3
	Right	24	85.7
	Ambidexter	0	0
Hearing	Good	22	78.6
	Moderate	5	17.9
	One good/one moderate	1	3.6
	Poor	0	0
Hearing device	Yes	1	3.6
	No	27	96.4
Cell phone handedness	Left	7	25.0
	Right	17	60.7
	Earphones	3	10.7

	Switch during conversation	1	3.6

Results of the healthy control group are summarized in Table 11. The earlier determined cutoff scores with regard to the norm data are a standard deviation (SD) of ≥1.5 for a score that is below average and an SD of ≥2 for a score that is deviant.

A more detailed description of the results per subtests on item level is provided in the next sections.

Table 11. Results of the healthy control group TeleTaalTest-NL.

Subtest (total score)	Mean	SD	Range
Comprehension screening (5)	4.93	0.26	4 - 5
Word finding (40)	37.57	3.06	28 - 40
• first ten items (20)	19.00	1.39	16 - 20
• additional ten items (20)	18.57	2.15	12 - 20
Phonological repetition (18)	17.75	0.44	17 - 18
Semantic selection (20)	19.50	0.92	17 - 20
Sentence completion (20)	18.86	1.21	16 - 20
TTT-NL final score (103)	98.61	3.98	84 - 103
• only first ten items word finding (83)	79.96	2.33	72 - 83
Cutoff scores	Below average		Deviant
Comprehension screening (5)	≤ 4.54		≤ 4.41
Word finding (40)	≤ 32.98		≤ 31.45
• first ten items (20)	≤ 16.92		≤ 16.22
• additional ten items (20)	≤ 15.35		≤ 14.27
Phonological repetition (18)	≤ 17.09		≤ 16.87
Semantic selection (20)	≤ 18.12		≤ 17.66
Sentence completion (20)	≤ 17.05		≤ 16.44
TTT-NL final score (103)	≤ 92.64		≤ 90.65
• only first ten items word finding (83)	≤ 76.47		≤ 75.30

Comprehension screening

The answers to all items were in line with the target answer, except for items 1 and 2, to which alternative answers have been given, see Table 12.

Table 12. Items comprehension screening to which alternative answers were given.

Comprehension screening				
Item	Question	Target answer	Alternative answers given	Rate of correctness
1.	*Is uw achternaam Jansen?* 'Is your surname Jansen?'	*Nee* 'No'	*Ja* 'Yes'[6] (1, correct)	100%
2.	*Bent u een kind?* 'Are you a child?'	*Nee* 'No'	*Ja* 'Yes' (2): *Ik ben een kind van mijn moeder.* 'I am my mother's child.' *Je bent een kind zo lang je ouders leven.* 'For as long as	92.9%

[6] The participant's surname indeed was Jansen, therefore, this answer was scored 'correct'.

			your parents are alive, you are a child.'	

Word finding

During administration of word finding, many alternative answers were given of which some have been scored correct, see Table 13.

Table 13. Items word finding to which incorrect or alternative answers were given.

Word finding				
Item	Question	Target word	Alternative answers given	Rate of correctness
3	*Hoe noem je dat als je elkaar spreekt via de telefoon?* 'What do you call it when you speak to each other via telephone?'	*bellen* 'to call'	*telefoneren* 'to telephone' (4, correct); *opbellen* 'to ring up' (1, correct); *appen* 'to whatsapp/text' (1); *gesprek* 'conversation' (1); *telefoongesprek* 'phone conversation' (3); *communicatie* 'communication' (1); *weet ik niet* 'I do not know' (1)	75%
4	*Wat hoor je op de radio?* 'What do you hear on the radio?'	*muziek* 'music'	*muziek onder anderen* 'music, amongst others' (1, correct); *stemgeluid* 'voice' (1, correct); *geluid* 'sound' (3, correct); *geluidsopnames* 'sound-recordings' (1, correct); *muziek of een verhaal* 'music or a story' (1, correct)	100%
5	*Bij een wedstrijd wint het ene team, wat doet het andere team?* 'At a game, one team wins; what does the other team do?'	*verliezen* 'to lose'	*verliest* 'loses' (4, correct); *gelijkspelen of verliezen* 'to draw or lose' (1, correct); *weet ik niet* 'I do not know' (1)	96.4%
7	*Wat doen mensen graag bij warm weer in de zee?* 'What do people like to do in the sea in hot weather?'	*zwemmen* 'to swim'	*zwemmen of baden* 'to swim or bathe' (1, correct)	100%
9	*Waar worden broden en taarten gemaakt?* 'Where are bread and cakes being made?'	*bakkerij* 'bakery'	*bakker* 'baker' (14, correct); *banketbakker* 'cake shop' (2, correct); *oven of bakkerij* 'oven or bakery' (1, correct); *oven* 'oven' (1)	96.4%
10	*Hoe noem je het wanneer een tuinman takken van een boom knipt?*	*snoeien* 'to prune'	*snoeien of kandelaberen* 'to prune or prune to the main branches' (1, correct); *knippen* 'to cut' (2); *hovenier* 'horticulturist' (1); *snijden* 'to	82.1%

	'What do you call it when a gardener cuts branches off a tree?'		cut' (1); *weet ik niet* 'I do not know' (1)	
11	*Een agent is onderdeel van de …* 'A police officer is part of the …'	*politie* 'police'	*politiekorps* 'police force' (3, correct); *controlerende macht* 'controlling power' (1); *maatschappij* 'society' (1); *weet ik niet* 'I do not know' (1)	89.3%
12	*Waar krijgen kinderen les?* 'Where do children have classes?'	*school* 'school'	*lokaal* 'classroom' (2, correct); *lokaal, school* 'classroom, school' (1, correct)	100%
13	*Wat doet een kind met poppen?* 'What does a child do with dolls?'	*spelen* 'to play'	*spelen, aankleden, meerdere mogelijkheden* 'to play, dress, multiple possibilities' (1)	96.4%
14	*Hoe noem je meerdere soldaten bij elkaar?* 'What do you call multiple soldiers together?'	*leger* 'army'	*korps* 'corps' (1, correct); *peloton* 'platoon' (1, correct); *army* 'army' (1); *militairen* 'military' (1); *weet ik niet* 'I do not know' (1)	89.3%
15	*Hoe heet bewegen op muziek ook wel?* 'What is moving to music also called?'	*dansen* 'to dance'	*dans* 'dance' (1); *weet ik niet* 'I do not know' (1)	92.3%
16	*Hoe heet het als je de waarheid niet vertelt?* 'What is it called when you do not tell the truth?'	*liegen* 'to lie'	*jokken* 'to fib' (1, correct); *meineed* 'perjury' (1); *leugen* 'lie' (1)	92.3%
17	*Hoe noem je het wanneer iemand eenden stukjes brood geeft?* 'What do you call it when someone gives ducks pieces of bread?'	*voeren/voederen (BE)* 'to feed'	*eenden/eendjes voeren* 'to feed ducks' (2, correct); *weet ik niet* 'I do not know' (1)	96.4%
18	*Waarin wonen een prins en prinses?* 'Where do a prince and princess live?'	*kasteel* 'castle'	*paleis* 'palace' (10, correct)	100%
20	*Wat doet iemand die mensen in trance brengt?* 'What does someone who entrances people do?'	*hypnotiseren* 'to hypnotise'	*hypnotiseur* 'hypnotist' (2); *emotioneel* 'emotional' (1); *hypnose* 'hypnosis' (1); *weet ik niet* 'I do not know' (4)	71.4%

Phonological repetition

Most items were successfully repeated, however, two items yielded alternative answers, see Table 14.

Table 14. Items phonological repetition to which incorrect answers have been given.

Phonological repetition			
Item	Target word	Alternative answers given	Rate of correctness
6	*prekamlédi*	*prekamlébi* (1); *prekomlédi* (1); *prekanlédi* (3); *pekanlédi* (1)	78.6%
8	*De kinderen fietsen naar huis.* 'The children cycle home.'	*De kinderen *hm hm* naar huis.* 'The children *hm hm* home.' (1)	96.4%

Semantic selection

Three items in the semantic selection task yielded alternative and incorrect answers, see Table 15.

Table 15. Items semantic selection to which incorrect answers have been given.

Semantic selection				
Item	Question	Target word	Answer	Rate of correctness
1	*boek zee krant* 'book sea newspaper'	*zee* 'sea'	*boek* 'book'[7] (2)	92.3%
4	*piano fluit koe* 'piano flute cow'	*fluit* 'flute'	Test leader spelled *fluit* 'flute', hence incorrect (1).	96.4%
9	*emmer brief gieter* 'bucket letter watering can'	*brief* 'letter'	*weet ik niet* 'I do not know' (1)	96.4%

Sentence completion

Despite there being various possible ways to finish the given sentences, some participants finished them incorrectly, see Table 16.

Table 16. Items sentence completion to which incorrect answers have been given.

Sentence completion			
Item	Question	Answers given	Rate of correctness
2	*De pan staat op …* 'The pan is on …'	Test leader spelled *pan* 'pan', hence incorrect (1).	96.4%
5	*De jongen gelooft dat …* 'The boy believes that …'	*weet ik niet* 'I do not know' (1)	96.4%
10	*Het meisje benadrukt dat …* 'The girl emphasizes that …'	*Het meisje benadrukte mijn moeder.* 'The girl emphasized my mother.' (1)	96.4%

Due to the nature of the test, some items were repeated when the participant could not clearly hear the item the first time. The test leader kept track of the number of repetitions in a separate column to see which answers needed repetition most, see Table 17.

[7] Participants misheard *krant* 'newspaper' for *strand* 'beach'.

Table 17. Number of repetitions per item.

Subtest	Item	Participant	Repetitions	Spelled out	Word	Interpretation
Comprehension screening	2. *Bent u een kind?* 'Are you a child?'	T_2 T_5 T_18	1 1 1	- - -	- - -	- - -
	3. *Is het nu winter?** 'Is it winter now?'	T_12 T_14 T_19 T_21	1 1 1 1	- - - -	- - - -	- - - -
Word finding	Practice item: *Waarin slaap je 's nachts?* 'Where do you sleep at night?'	T_8 T_12 T_17	1 1 1	- - -	*waarin* 'where' - *waarin* 'where'	*waarom* 'why' - -
	1. Er zitten 60 minuten in een … 'There are 60 minutes in an ...'	T_14	1	-	-	-
	9. *Waar worden broden en taarten gemaakt?* 'Where are bread and cakes being made?'	T_14	1	-	-	-
	13. *Wat doet een kind met poppen?* 'What does a child do with dolls?'	T_2 T_12 T_19 T_23	1 1 1 1	- - - -	- - - -	- - - -
	15. *Hoe heet bewegen op muziek ook wel?* 'What is moving to music also called?'	T_12	1	-	-	-
	20. *Wat doet iemand die mensen in trance brengt?* 'What does someone who entrances people do?'	T_11 T_14 T_16 T_18 T_28 T_29	1 1 1 1 1 1	- - - - - -	- - - - - -	- - - - - -
Phonological repetition	Practice item: *Legende* 'legend'	T_7 T_14	1 3	- -	*legende* 'legend' *legende* 'legend'	- *lezjende*
	8. *De kinderen fietsen naar huis.* 'The children cycle home.'	T_12	1	-	*fietsen naar* 'cycle'	-

Semantic selection	Practice item: *Vlieg, schip, kever* 'fly, ship, beatle'	T_1 T_2 T_14 T_21	1 1 4 1	- - - -	*vlieg* 'fly' *vlieg* 'fly' *vlieg/kever* 'fly/beatle' *kever* 'beatle'	*vis* 'fish' - *vlies* 'film'*/stever, steven* -
	1. boek, zee, krant 'book, sea, newspaper'	T_3 T_7 T_10 T_18 T_23 T_25	1 1 1 1 1 1	- - - - - -	*krant* 'newspaper' *krant* 'newspaper' *krant* 'newspaper' *zee* 'sea' *zee* 'sea' *zee* 'sea'	- - - - - -
	2. *sla, wortel, huis* 'cabbage, carrot, house'	T_14	1	-	*huis* 'house'	*rijst* 'rice'
	3. kerk, hond, paard 'church, dog, horse'	T_3 T_5 T_7 T_9 T_15 T_18 T_24	1 2 1 1 1 1 1	- - - - - - -	*hond* 'dog' *hond* 'dog' *hond* 'dog' *hond* 'dog' *paard* 'horse' *hond* 'dog' *kerk* 'church'	- - - - - - -
	4. piano, fluit, koe 'piano, flute, cow'	T_22 T_28 T_30	3 1 1	1 - -	*fluit* 'flute' *fluit* 'flute' *fluit* 'flute'	- - -
	5. zomer, schaar, winter 'summer, scissors, winter'	T_14	1	-	-	-
	6. *peer, appel, land* 'pear, apple, land'	T_21	1	-	*land* 'land'	-
Sentence completion	Practice item 2: *Iedere maand …* 'Every month ...'	T_7 T_18 T_30	1 1 1	- - -	- - -	- - -
	1. Hij valt van … 'He falls from ...'	T_11 T_14 T_19	1 1 1	- - -	- - -	- - -
	2. *De pan staat op …* 'The pan is on ....'	T_5 T_10 T_12	5 1 2	2 - -	*pan* 'pan' *pan* 'pan' -	*stand* 'score' - -

		T_29 T_31	3 1	- -	- *pan* ‘pan’	- *tand* ‘tooth’
	3. Ik snijd met … ‘I cut with …’	T_12	1	-	-	-
	4. De vrouw probeerde … ‘The woman tried …’	T_29	1	-	-	-
	8. *Zij belde niet omdat …* ‘She did not call because …’	T_14	2	-	*niet* ‘not’	*u* ‘you’
	9. *Toen de vaas brak …* ‘When the vase broke …’	T_12 T_15 T_23	1 1 1	- - -	- *vaas brak* ‘vase broke’ -	- - -
	10. *Het meisje benadrukt dat …* ‘The girl emphasizes that …’	T_19	1	-	-	-

Non-parametric tests (Kruskal-Wallis and Mann-Whitney U) were run to analyze TTT-NL subtest and final scores from age, education in years, educational level on the Verhage-scale, gender, handedness, hearing, hearing device, and the way the telephone was held by the participant, see Table 18 and Table 19 for TTT-NL (sub)scores according to age, education in years and education Verhage.

TTT-NL scores: age

A Mann-Whitney U test revealed a significant effect of age for the sentence completion task, where participants between 18 and 54 years old (Mdn = 20.00) scored significantly higher than participants between 55 and 75 years old (*Mdn* = 18.00), $U = 43.50$, $z = -2.48$, $p < .05$, $r = -.47$.

TTT-NL scores: education in years

There was a significant effect for education in years in the semantic selection task (> 12 years: *Mdn* = 20.00, ≤ 12 years: Mdn = 18.00), $U = 17.00$, $z = -2.69$, $p < .01$, $r = -.51$, the total score for the TTT-NL with one set of word finding items (> 12 years: *Mdn* = 81.00, ≤ 12 years: *Mdn* = 76.50), $U = 8.50$, $z = -2.66$, $p < .05$, $r = -.50$, and the total score for the TTT-NL (> 12 years: *Mdn* = 100.00, ≤ 12 years: *Mdn* = 96.50), $U = 18.00$, $z = -1.99$, $p < .05$, $r = -.38$. In every case, participants with more than twelve years of education outperformed participants with less than or equal to twelve years of education.

TTT-NL scores: education Verhage

A Kruskal-Wallis test showed a significant effect of education on the Verhage-scale for the total score of the second set of word finding items, $H(3) = 8.48$, $p < .05$. Participants classified level 7 (*Mdn* = 20.00) scored significantly higher than participants classified level 6 (*Mdn* = 20.00), 5 (*Mdn* = 18.00), or 4 (*Mdn* = 13.00). Similar significant effects of education on the Verhage-scale were found for phonological repetition (7: *Mdn* = 18.00, 6: *Mdn* = 18.00, 5: *Mdn* = 17.50, 4: *Mdn* = 17.00), $H(3) = 7.89$, $p < .05$, semantic selection (7: *Mdn* = 20.00, 6: *Mdn* = 20.00, 5: *Mdn* = 19.00, 4: *Mdn* = 18.00), $H(3) = 11.14$, $p < .05$, the total score for the TTT-NL with one set of word finding items (7: *Mdn* = 81.00, 6: *Mdn* = 80.00, 5: *Mdn* = 79.50, 4: *Mdn* = 74.00), $H(3) = 8.11$, $p < .05$, and the total score for the TTT-NL (7: *Mdn* = 101.00, 6: *Mdn* = 99.00, 5: *Mdn* = 97.50, 4: *Mdn* = 87.00), $H(3) = 9.04$, $p < .05$.

TTT-NL scores: hearing

There was a significant effect of hearing on the word finding total (wf tot.) 2 scores, $H(2) = 8.95$, $p < .05$. Participants with good hearing (*Mdn*: 20.00) outperformed participants with moderate (*Mdn*: 16.00), or one good and the other ear moderate (good/moderate) hearing (*Mdn*: 16.00). Similar significant effects of hearing were found for the word finding total score, $H(2) = 7.99$, $p < .05$, the total score for the TTT-NL with one set of word finding items, $H(2) = 6.49$, $p < .05$, and the total score for the TTT-NL, $H(2) = 7.39$, $p < .05$. The group who outperformed the other groups differed per subtest/total score. In word finding total and the total score for the TTT-NL, participants with good hearing (wf tot.: *Mdn* = 40.00, TTT-NL: *Mdn* = 100.00) outperformed participants with good/moderate (wf tot.: *Mdn* = 36.00, TTT-NL: *Mdn* = 97.00), and moderate hearing (wf tot.: *Mdn* = 32.00, TTT-NL: *Mdn* = 95.00). In the total score for the TTT-NL with one set of word finding items, participants with good/moderate hearing (*Mdn* = 81.00) outperformed participants with good (*Mdn* = 81.00), and moderate hearing (*Mdn* = 79.00).

No significant effects were found for gender, handedness, hearing device, or the way in which the participant held the telephone against his ear.

Table 18. Kruskal-Wallis/Mann-Whitney U test results for demographic variables and TTT-NL subtest and final scores.

	Age in years	Education in years	Education Verhage	Gender	Handedness	Hearing	Hearing device	Telephone
Comprehension screening	.246	.141	.948	.122	.556	.753	.782	.719
Word finding total 1	.684	.112	.181	.441	.472	.058	.433	.887
Word finding total 2	.377	.823	.037*	.958	.390	.011*	.481	.382
Word finding total all	.458	.557	.099	.753	.370	.018*	.845	.462
Repetition total	.273	.221	.048*	.830	1.000	.619	.564	.484
Semantic selection	.828	.007**	.011*	.692	.795	.624	.567	.285
Sentence completion	.013*	.104	.133	.771	.533	.194	.133	.489
TTT-NL total score	.225	.046*	.029*	.889	.207	.025*	.573	.479
TTT-NL total score (1x word finding)	.111	.008**	.044*	.554	.711	.039*	.656	.684

Note. * p < .05; ** p <,01; *** p < .001; n = 28; two-tailed.

Table 19. TTT-NL (sub)scores according to age, education in years, and education Verhage.

Subtest (total score)	Mean				SD				Range			
	Age		Education in years		Age		Education in years		Age		Education in years	
	< 55	≥ 55	≤ 12	> 12	< 55	≥ 55	≤ 12	> 12	< 55	≥ 55	≤ 12	> 12
Comprehension screening (5)	4.88	5.00	4.75	4.96	0.33	0.00	0.50	0.20	4 - 5	5 - 5	4 - 5	4 - 5
Word finding (40)	38.06	36.82	36.00	37.83	2.41	3.87	5.42	2.58	32 - 40	28 - 40	28 - 40	32 - 40
• first ten items (20)	19.06	18.91	18.00	19.17	1.44	1.38	1.63	1.31	16 - 20	16 - 20	16 - 20	16 - 20
• additional ten items (20)	19.00	17.91	18.00	18.67	1.50	2.84	4.00	1.81	16 - 20	12 - 20	12 - 20	14 - 20
Phonological repetition (18)	17.82	17.64	17.50	17.79	0.39	0.51	0.58	0.42	17 - 18	17 - 18	17 - 18	17 - 18
Semantic selection (20)	19.53	19.45	**18.25**	**19.71**	0.94	0.93	**1.26**	**0.69**	17 - 20	18 - 20	**17 - 20**	**18 - 20**
Sentence completion (20)	**19.29**	**18.18**	17.75	19.04	**1.05**	**1.17**	1.71	1.04	**17 - 20**	**16 - 20**	16 - 20	17 - 20
TTT-NL final score (103)	99.59	97.09	**94.25**	**99.33**	2.48	5.36	**7.04**	**2.87**	95 - 103	84 - 101	**84 - 100**	**90 - 103**
• only first ten items word finding (83)	80.47	79.18	**76.25**	**80.58**	1.94	2.75	**3.30**	**1.47**	76 - 83	72 - 81	**72 - 80**	**76 - 83**

Subtest (total score)	Mean				SD				Range			
	Education Verhage				Education Verhage				Education Verhage			
	4	5	6	7	4	5	6	7	4	5	6	7
Comprehension screening (5)	5.00	5.00	4.91	4.92	0.00	0.00	0.30	0.28	5 - 5	5 - 5	4 - 5	4 - 5
Word finding (40)	30.00	36.00	38.00	38.62	2.83	5.66	2.24	1.50	28 - 32	32 - 40	34 - 40	36 - 40
• first ten items (20)	17.00	18.00	19.27	19.23	1.41	2.83	1.35	1.01	16 - 18	16 - 20	16 - 20	18 - 20
• additional ten items (20)	**13.00**	**18.00**	**18.73**	**19.38**	**1.41**	**2.83**	**1.56**	**1.26**	**12 - 14**	**16 - 20**	**16 - 20**	**16 - 20**
Phonological repetition (18)	**17.00**	**17.50**	**17.91**	**17.77**	**0.00**	**0.71**	**0.30**	**0.44**	**17 - 17**	**17 - 18**	**17 - 18**	**17 - 18**
Semantic selection (20)	**18.00**	**19.00**	**19.27**	**20.00**	**0.00**	**1.41**	**1.10**	**0.00**	**18 - 18**	**18 - 20**	**17 - 20**	**20 - 20**
Sentence completion (20)	17.00	20.00	18.82	19.00	1.41	0.00	1.08	1.16	16 - 18	20 - 20	17 - 20	17 - 20
TTT-NL final score (103)	**87.00**	**97.50**	**98.91**	**100.31**	**4.24**	**3.54**	**2.17**	**1.80**	**84 - 90**	**95 - 100**	**96 - 103**	**97 - 103**
• only first ten items word finding (83)	**74.00**	**79.50**	**80.18**	**80.77**	**2.83**	**0.71**	**1.83**	**1.42**	**72 - 76**	**79 - 80**	**76 - 83**	**77 - 83**

Spearman correlation tests were run on TTT-NL (sub)scores, age-matched BNT norm data, and DIMA (sub)scores, see Appendix A, to assess the external validity of the TTT-NL, see Table 20. There was a significant correlation between semantic selection and DIMA SEM A, ($r_s(27) = .149$, $p < .05$. No significant correlations were found between the word finding scores and the age-matched BNT norm data, or the between the remaining TTT-NL scores and the DIMA (sub)scores.

Table 20. Correlation table BNT norms, DIMA control data and TTT-NL subtest and final scores.

	BNT	DIMA PHON total	DIMA SEM A	DIMA SYN B	DIMA total (only SEM A)
Word finding total 1	.182	-	-	-	-
Word finding total 2	.321	-	-	-	-
Word finding total all	.284	-	-	-	-
Repetition total	-	.460	-	-	-
Semantic selection	-	-	.021*	-	-
Sentence completion	-	-	-	.250	-
TTT-NL total score	-	-	-	-	.653
TTT-NL total score (1x word finding)	-	-	-	-	.447

Note. * p < .05; ** p <,01; *** p < .001; n = 28; two-tailed.

Chapter 4: Outcomes

4.1 Discussion

The present study aimed to improve the existing version of the TeleTaalTest-NL, to assess the improved version in a healthy control group, and to compare the TTT-NL subtests with DIMA subtests in order to evaluate if the test is a valid alternative for in-person assessment. In doing so, the following research questions were addressed:

(1) Is the TeleTaalTest-NL (Satoer et al., 2020a) a valid tele-language test when administered in a healthy control group?

(2) Do the TeleTaalTest-NL (Satoer et al., 2020a) and its subtests correlate with the DIMA (Satoer et al., 2019) and its subtests?

Assessment of a healthy control group

The new version of the TTT-NL - with an improved comprehension screening, word finding, verbal repetition, and sentence completion - was administered in a healthy control group twice; once within the pilot study and once within the norm study. The rate of correctness of all items for the norm study, which had to be 80% or above, had been assessed in the pilot study. In both studies, the control groups performed well as only few mistakes were made. Effects of age, education in years, education on the Verhage-scale, and hearing were found for word finding total 2 (education Verhage, hearing), word finding total all (hearing), verbal repetition (education Verhage), semantic selection (education in years, education Verhage), sentence completion (age), and TTT-NL final scores (education in years, education Verhage, hearing). This is in agreement with previous studies on language assessment where similar effects of age (the higher the age, the lower the scores) and education (the higher the education, the higher the scores) were found (De Witte et al., 2015; Satoer et al., 2019). In general participants with good hearing performed better than those with good/moderate or moderate hearing during telephone assessment, which also is consistent with previous research (Newkirk et al., 2004; Roccaforte et al., 1992).

External validity

The correlation between in-person assessment and tele-assessment of the TTT-NL had not yet been assessed in previous research. A significant correlation was found between semantic selection and SEM A, however, the remaining TTT-NL (sub)scores did not significantly correlate to corresponding DIMA and/or BNT (sub)scores.

Previous research on the correlation between telephone and face-to-face assessment of language and cognitive abilities did find a correlation between the two methods of assessment. In general, there is little to no difference between the results of assessment with either one of these methods (Mitsis et al., 2014; Pendlebury et al., 2013; Van den Berg et al., 2012) and sometimes tele-assessment is even more sensitive than in-person assessment (De Witte et al., 2019), which shows that tele-assessment is a valid alternative for face-to-face assessment. However, in some cases, it is more difficult to detect subtle impairments with tele-assessment, which often is a consequence of a relatively small sample size (De Witte et al., 2019; Pendlebury et al., 2013; Van den Berg et al., 2012; Zietemann et al., 2017). Translated to the current study, of which the sample size for the healthy controls tested with the TTT-NL is very small, this means that it is plausible that the sample size is the cause for the absence of significant correlations between telephone and in-person assessment. Furthermore, where previous studies applied both methods of assessment within one group (De Witte et al., 2019; Mitsis et al., 2014; Pendlebury et al., 2013; Van den Berg et al., 2012; Zietemann et al., 2017), the current study only applied tele-testing in a group of 28 healthy participants as opposed to a sample of 214 healthy participants for

in-person assessment of the DIMA. The problem could have been by-passed by (1) also assessing the DIMA within the same small-sized control group; or (2) expanding the current TTT-NL healthy control sample to a group size that is a faithful reflection of society, which makes it comparable to the group in which the DIMA was assessed.

Alternative explanations for the absence of significant correlations are the use of age-matched BNT mean norm scores instead of BNT scores obtained by participants themselves in word finding, the large difference between the number of items for phonological repetition of the TTT-NL (10 items) and the DIMA (40 items), and the obtained ceiling levels and small standard deviations for the TTT-NL, which leave limited space for correlations. The latter could be remedied by expanding the sample size or by increasing statistical power by performing one-sided instead of two-sided tests (Brock, 2019).

Deviant responses and repetitions

Not all items of the TTT-NL yielded target answers. A too wide scope of the question led to ambiguity, unclarity about the word class of the target word, and/or either too broad or more specific responses than initially aimed for. Some items were ambiguous, such as item 2 of the comprehension screening, where *kind* 'child' can be interpreted in two ways, with reference to age (≥ 18 years) or the presence of parents (you are a child for as long as your parents are alive). Substitution would solve an ambiguity problem. Item 2 could for instance be substituted for an item of similar length, such as *Bent u twaalf jaar?* 'Are you twelve years old?', by-passing the ambiguity of the word *kind* 'child', yet testing the same question with the similar target answer *nee* 'no'. Unclarity about the word class of the target word especially was a problem in word finding. Item 3 *Hoe noem je dat als je elkaar spreekt via de telefoon?* 'What do you call it when you speak to each other via telephone?', for instance, could have both the verb *bellen* 'to call' (target word) or the noun *telefoongesprek* 'phone conversation' as the target word. Some answers were of the right word class, however, answers were either too broad, such as *controlerende macht* 'controlling power' or *maatschappij* 'society' instead of the target word *politie* 'police' for item 11, or too specific, such as *lokaal* 'classroom' instead of the target word *school* 'school' for item 12, both within word finding. Unclarities about word class and answers of varying scopes can be resolved by specifying the question, such as the use of *agenten* 'police officers' instead of the singular form *agent* 'police officer' to emphasize that you are referring to the police as a group, or the replacement of *onderdeel van* 'part of' with *lid van* 'member of' to direct the patient toward the organization that a police officer is part of (item 11 word finding).

Due to a combination of low frequency words (especially in word finding) and time pressure in most subtests and/or the feeling of being subjected to pressure or stress in general, participants gave incorrect or alternative answers. *Snoeien* 'to prune' and *hypnotiseren* 'to hypnotize' (items 10 and 20 word finding) are both low frequency words to which some participants gave various (often incorrect) answers, amongst others *weet ik niet* 'I do not know' and *knippen* 'to cut' for *snoeien* 'to prune'. This is in line with previous research, as low frequency words require more time for lexical access in tasks such as naming (Forster & Chambers, 1973; Oldfield & Wingfield, 1965) and are more often subject to the tip-of-the-tongue state[8] (Gollan et al., 2014), often leading to the naming a word that is similar to the target word (Brown & McNeill, 1966). Furthermore, item 15 of word finding *dansen* 'to dance' has a relatively high frequency, which means that it should be easier to retrieve the target word. This line of reasoning is confirmed by the high rate of correctness, however, one participant still answered with *weet ik niet* 'I do not know'. Word finding does apply a timeframe of four seconds, which means that time pressure most likely is the cause for the incorrect answer.

Finally, participants often misheard words or speech sounds due to the quality of the connection, the slight southern accent of the test leader, and/or the lack of visuals. Overall, the quality of the connection was good, however, it sometimes varied due to multiple factors, such as whether a landline or mobile telephone was used. This resulted in distortion or the loss of speech sounds, which made

[8] The tip-of-the-tongue state is the feeling that you temporarily cannot recall a familiar word but that you are able to retrieve a word of similar form and/or meaning (Brown & McNeill, 1966; Schwartz & Metcalfe, 2011).

assessment of the TTT-NL harder for both parties. Additionally, approximately one third of the participants pointed out that the test leader's slight southern accent contributed to mishearing words or speech sounds, as in for example *vlieg* 'fly' (practice item semantic selection), where the *g* (IPA: [x]) could barely be heard. Third, some participants indicated that, for example during phonological repetition, it was more challenging for them to distinguish certain speech sounds and/or words (e.g. fricatives, such as in *hond* 'dog') because they could not see the articulatory movements of the test leader, which normally are supportive during a conversation or a task such as phonological repetition (Grant & Seitz, 2000; Sumby & Pollack, 1954). This was especially the case for speech sounds with a closely related manner of articulation (e.g. nasals *n* and *m* and plosives *p* and *d*) or tongue position (e.g. *ɑ* and *ɔ*, both back vowels). The problem mostly became evident in tasks that consist of words and/or sentences with no or very little context to derive the right word from (e.g. phonological repetition, semantic selection, and sentence completion) or of non-existing words (e.g. phonological repetition). Nonwords are more difficult to repeat as they have not been heard before and are not in the lexicon (Sierpowska et al., 2017). Hence, this increases the intelligibility problem and could explain the relatively high number of incorrect repetitions of *prekanlédi*. De Witte and colleagues (2019) reached the same conclusion and decided to leave nonwords out altogether as more than the verbal in- and output provided by tele-testing is needed to reliably assess nonword repetition. The nonwords within this study - especially *prekanlédi* - should therefore be reconsidered and changed or left out.

The earlier mentioned causes for deviant responses also led to repetitions, with low frequency words (e.g. item 20 word finding) and unintelligible speech sounds (e.g. item 3 semantic selection and item 2 sentence completion) yielding the most repetitions. At worst, items had to be spelled out (i.e. *pan* 'pan' in item 2 sentence completion) as multiple repetitions did not provide clarification.

4.2 Clinical implications

The TTT-NL enables patients who are not able to come to the outpatient's clinic for aphasia diagnostics to be assessed from a distance, through which possible language disorders can still be observed timely. Although in the present study, the test was only administered in a healthy control group, the TTT-NL is expected to be applicable for signaling (mild) aphasia in a population with a neurological disorder by telephone as each of the subtests addresses other symptoms that can be present in the different aphasia subtypes (see section 1.1).

Nevertheless, certain groups who suffer from a more severe type of aphasia affecting comprehension, cannot be assessed by telephone. These patients with a so-called comprehension deficit are filtered out by the comprehension screening, which was designed to assess auditive comprehension. The screening is sensitive for patients with global aphasia (all modalities are impaired), Wernicke's aphasia (impaired auditive input lexicon), transcortical sensory aphasia (deficit between the auditive input lexicon and the semantic system), and in case of additional comprehension deficits, mixed dynamic aphasia[9], timely ending assessment of the TTT-NL for this population.

Patients with intact comprehension proceed to the subtests of the TTT-NL, starting with word finding. This task was designed to assess naming via the phonological in- and output route, which makes this subtest sensitive for patients with anomic aphasia (synonyms/descriptions of the target word or empty speech due to a deficit between the semantic system and the phonological output lexicon) and possibly conduction aphasia (phonemic paraphasias and conduite d'approche when naming the target word due to a deficit in the phonological output lexicon and phonological coding) although phonemic paraphasias are mostly present in repetition.

Second, the phonological repetition subtest was designed to assess verbal repetition via the phonological in- and output route (words, compounds, and sentences) and the direct route between the auditive analysis system and phonological coding (nonwords). Difficulties in repetition can be an

[9] Besides sentence completion as the main subtest, all subtests can be sensitive for mixed dynamic aphasia depending on the patient's additional impairments.

indication of Broca's aphasia (telegram-style and impaired repetition of function words and sentences due to a grammatical coding deficit and possibly an articulatory deficit) or conduction aphasia.

Semantic selection was designed to assess semantic judgment via the auditive in- and output route, making this a more challenging and sensitive subtest for patients with conduction aphasia.

The final subtest of the TTT-NL was designed to assess spontaneous speech in context. Sentence completion combines all elements of previous subtests, which makes it possible to detect symptoms of anomic aphasia, Broca's aphasia (additional problems in the initiation of speech), conduction aphasia (although phonemic paraphasias are mostly present in repetition), transcortical motoric aphasia (non-fluent speech and problems in the initiation of speech due to a grammatical coding deficit), mixed dynamic aphasia (trouble to select a response due to a verbal response generation deficit), and pure dynamic aphasia (trouble to select a response due to a verbal response generation deficit).

Hence, the TTT-NL has the potential to cover many language levels, facilitating the observation of possible symptoms of various aphasia subtypes. In practice, patient scores can be compared to cutoff scores, which were determined based on age and education. There is no need for additional materials besides a (stable) telephone connection, making this a method of assessment that is highly accessible to everyone.

4.3 Limitations and future directions

The findings of the present study have to be seen in the light of some limitations, which were partly caused by the COVID-19 pandemic.

A few participants gave multiple 'options' in their answer instead of providing only one (target) word in word finding, such as *muziek, onder anderen* 'music, amongst others' (item 4), making scoring more ambiguous and subject to larger inter-rater variability. Multiple-word answers can be reduced by a question to clarify the response, along the lines of *En als je het in één woord moet zeggen?* 'And if you had to say it in one word?'. Clear rules and permitted synonyms should be included in the manual, facilitating scoring with one accord.

Second, as mentioned in section 4.1, the current sample size is small, which most likely led to the absence of correlations between the TTT-NL and in-person assessment. Additionally, the BNT scores used for assessing the external validity of the TTT-NL were not obtained by the healthy population. Picture naming involves visuals and no online version of the test exists yet. Longer preparation time could have resulted in finding a technical solution (i.e. administration via Microsoft Teams, Skype, or a different medium), however, time was limited, which is why personal, and therefore more reliable, BNT scores are missing. For more reliable and complete tele-diagnostics, future research could focus on the development of digital tools, such as the already available, digital version of the Token Test (Satoer et al., 2020b).

Third, there are a couple more general limitations. External factors, such as fatigue and environmental noise can never be fully controlled for. This conclusion is similar to the one that De Witte and colleagues (2019) reached. A limitation as such is due to the nature of the test, which cannot easily be remedied but should be kept at the back of one's mind when administering the test and interpreting the obtained results. One of the external factors that can be controlled for to some extent is the quality of the connection. During one of the assessments, both the test leader's and the participant's phone was on speaker, creating an echo that caused intelligibility problems for both parties, which made the test results unreliable[10]. In future use of the TTT-NL, only the test leader should put his phone on speaker. The use of two landline phones (the test leader always used one) also contributed to a stable connection. Additionally, possible alternatives for audio recording the conversation (e.g. calling with headphones) should be explored as approximately half of the participants expressed the connection was of better quality when the test leader's phone was not on speaker.

[10] This participant was excluded from analysis.

As mentioned in section 4.1, during the administration of the TTT-NL, items were repeated multiple times and sometimes spelled out. An additional rule that permits only one repetition per item should be entered in the manual as it is hard to discover the reason for the demanded repetition (e.g. delaying tactics).

As opposed to the studies by Ruijs (2018) and Van Dijk (2018), the present study contains only one version of the TTT-NL. In order to control for a learning-effect, a parallel version is essential. Due to the late change of topic, there was not enough time to develop a second version of the test and to administer both versions. As having two versions of the TTT-NL is desirable, future research could be aimed at the development of a parallel version of the current test.

Before assessment, no Mini Mental State Examination (MMSE) (Kok & Verhey, 2002) was administered in the healthy population, meaning that cognitive decline after the age of 54 (see section 2.1) cannot be ruled out. Not all questions of the MMSE are suitable for telephone assessment. The t-MoCA (Pendlebury et al., 2013) would have been an appropriate alternative (Hoops et al., 2009; Webb et al., 2014), however, due to a lack of time this option was not considered. Therefore, in future research, participants included in the healthy control group of age 55 and above should be tested with either the MMSE or t-MoCA to control for cognitive decline.

Referring back to the clinical applications of the TTT-NL, sometimes patients are too severely aphasic to be tested with the tele-test (e.g. global aphasia). Video calling could be an alternative as this increases the input for both parties, also creating possibilities for the assessment of tests of a different nature, such as tests involving visual input (e.g. picture naming). Nevertheless, face-to-face assessment should remain available for this group of patients. Furthermore, the TTT-NL does not include subtests that assess the modalities writing, reading, or picture naming; neither does it provide visual information about the articulation of the participant, meaning that it is harder, if not impossible, to signal speech apraxia, dysarthria, and surface dyslexia by phone and to distinguish between “classical” aphasia syndromes. Hence, the TTT-NL is not a replacement for face-to-face language testing and existing language tests but it is a supplement and useful tool in situations such as the COVID-19 pandemic.

The current version of the TTT-NL is based on subtests that are frequently used in (mild) aphasia tests, such as the verbal naming test (Yochim et al., 2015) and the WAB (Kertesz, 2007), meaning that, despite the mentioned limitations, this does not change anything with reference to the suitability of the TTT-NL subtests to signal language impairments. Further research of a larger sample size that takes the current limitations into account is needed.

Chapter 5: Conclusions

The goal of the present study was to improve and supplement the healthy control data of the current version of the TTT-NL, and to assess its external validity by comparing healthy control scores of the TTT-NL to healthy control scores of the DIMA and BNT norm data. Improvement of the TTT-NL was successful as participants within the pilot and norm study performed at ceiling level, meaning that the comprehension screening and subtests of the TTT-NL proved to be effective and feasible. Furthermore, patient scores can now be compared to cutoff scores based on age and education. Only semantic selection significantly correlated with DIMA norm data, which is most likely due to the small sample size.

Due to the absence of significant correlations between the TTT-NL and the DIMA, there is no evidence yet that indicates the TTT-NL is a reliable replacement of in-person assessment. Therefore, more research of a larger sample size is needed to further investigate the relationship between tele-testing and in-person language assessment.

References

Audacity Team (2020). *Audacity®* (Version 2.4.2). Distributed under the terms of the GNU General Public License. https://www.audacityteam.org/download/

Bastiaanse, R. (2011). *Afasie.* Houten, The Netherlands: Bohn Stafleu van Loghum.

Brock, J. (2019, September 16). *5 Tips for dealing with non-significant results: It might look like a failure, but don't let go just yet.* Nature Index. https://www.natureindex.com/news-blog/top-tips-for-dealing-with-non-significant-null-results#:~:text=The%20problem%20with%20a%20non,data%20are%20inconclusive%20either%20way.

Brown, R., & McNeill, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior, 5*(4), 325-337.

Castanho, T. C., Amorim, L., Zihl, J., Palha, J. A., Sousa, N., & Santos, N. C. (2014). Telephone-based screening tools for mild cognitive impairment and dementia in aging studies: A review of validated instruments. *Frontiers in Aging Neuroscience, 6*, 1-17.

De Renzi, E., & Faglioni, P. (1978). Normative data and screening power of a shortened version of the Token Test. *Cortex, 14*(1), 41-49.

De Witte, E., Satoer, D., Robert, E., Colle, H., Verheyen, S., Visch-Brink, E., & Mariën, P. (2015). The Dutch Linguistic Intraoperative Protocol: A valid linguistic approach to awake brain surgery. *Brain & Language, 140*, 35-48.

De Witte, E., Piai, V., Kurteff, G., Cai, R., Mariën, P., Dronkers, N., Chang, E., & Berger, M. (2019). A valid alternative for in-person language assessments in brain tumor patients: feasibility and validity measures of the new TeleLanguage test. *Neuro-Oncology Practice, 6*(2), 93-102.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12*(6), 627-635.

Gollan, T. H., Ferreira, V. S., Cera, C., & Flett, S. (2014). Translation-priming effects on tip-of-the-tongue states. *Language & Cognitive Processes, 29*(3), 278-288.

Goodglass, H., & Kaplan, E (1972). *The assessment of aphasia and related disorders.* Philadelphia, PA: Lea and Fibiger.

Goodglass, K., Kaplan, E., & Barresi, B. (2000). *Boston Diagnostic Aphasia Examination (BDAE-3)* (3rd ed.). Philadelphia, PA: Lea & Febiger.

Graetz, P., De Bleser, R., & Willmes, K. (1992). *Akense Afasie Test*. Amsterdam, The Netherlands: Hogrefe.

Grant, K. W., & Seitz, P. (2000). The use of visible speech cues for improving auditory detection of spoken sentences. *The Journal of the Acoustical Society of America, 108*(3), 1197-1208.

Grossman, M., & Irwin, D. J. (2018). Primary Progressive Aphasia and stroke aphasia. *Behavioral Neurology and Psychiatry, 24*(3), 745-767.

Hallowell, B. (2017). *Aphasia and other acquired neurogenic language disorders: A guide for clinical excellence.* San Diego, CA: Plural Publishing Inc.

Hersenz. (n.d.). *Feiten en cijfers over niet-aangeboren hersenletsel*. Hersenz. https://www.hersenz.nl/hersenletsel/feiten-cijfers-niet-aangeboren-hersenletsel

Hoops, S., Nazem, S., Siderowf, A. D., Duda, J. E., Xie, S. X., Stern, M. B., & Weintraub, D. (2009). Validity of the MoCA and MMSE in the detection of MCI and dementia in Parkinson disease. *Neurology, 73*(21), 1738-1745.

IBM Corp. Released 2020. IBM SPSS Statistics for Windows, Version 27.0. Armonk, NY: IBM Corp.

Incekara, F., Satoer, D., Visch-Brink, E., Vincent, A., & Smits, M. (2019). Changes in white matter tract microarchitecture associated with cognitive deficits in patients with presumed low-grade glioma. *Journal of Neurosurgery, 130*(5), 1538-1546.

Kaplan, E., Goodglass, H., & Weintrab, S. (1983). *The Boston Naming Test*. Philadelphia, PA: Lea & Febiger.

Kaplan, E., Goodglass, H., & Weintrab, S. (2001). *Boston Naming Test* (2nd ed.). Philadelphia, PA: Lippincott, Williams & Wilkins.

Kertesz, A. (2007). *Western Aphasia Battery Revised. Examiner's Manual.* San Antonio, TX: Harcourt Assessment Inc..

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods, 42*(3), 643-650.

Kok, R., & Verhey, F. (2002). *Dutch translation of the Mini Mental State Examination* (Folstein et al., 1975). Utrecht, The Netherlands: Altrecht, GGZ.

Marchina, S., Zhu, L. L., Norton, A., Zipse, L., Wan, C. Y., & Schlaug, G. (2011). Impairment of speech production predicted by lesion load of the left arcuate fasciculus. *Stroke, 42*(8), 2251-2256.

Mitsis, E. M., Jacobs, D., Luo, X., Andrews, H., Andrews, K., & Sano, M. (2010). Evaluating cognition in an elderly cohort via telephone assessment. *International Journal of Geriatric Psychiatry, 25*(5), 531-539.

Newkirk, L. A., Kim, J. M., Thompson, J. M., Tinklenberg, J. R., Yesavage, J. A., & Taylor, J. L. (2004). Validation of a 26-point telephone version of the Mini-Mental State Examination. *Journal of Geriatric Psychiatry and Neurology, 17*(2), 81-87.

Nichelli, P. (2016). Chapter 23: Consciousness and Aphasia. In S. Laureys, O. Gosseries, & G. Tononi (Eds.), *The neurology of consciousness: Cognitive neuroscience and neuropathology* (2nd ed.) (pp. 379-391). Amsterdam, The Netherlands: Elsevier Ltd.

Nouwens, F., Dippel, D. W. J., Visch-Brink, E. G., & de Lau, L. M. L. (2013). Behandeling van afasie door een beroerte. *Tijdschrift voor Neurologie & Neurochirurgie, 114*(2), 52-59.

Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal in Experimental Psychology, 17*(4), 273-281.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113.

Pendlebury, S. T., Welch, S. J. V., Cuthbertson, F. C., Mariz, J., Mehta, Z., & Rothwell, P. M. (2013). Telephone assessment of cognition after TIA and stroke: TICSm and telephone MoCA vs face-to-face MoCA and neuropsychological battery. *Stroke, 44*(1), 227-229.

Prete, G., D'Anselmo, A., Brancucci, A., & Tommasi, L. (2018). Evidence of Right Ear Advantage in the absence of auditory targets. *Scientific Reports, 8*, 15569.

Robinson, G., Shallice, T., & Cipolotti, L. (2006). Dynamic aphasia in progressive supranuclear palsy: A deficit in generating a fluent sequence of novel thought. *Neuropsychologia, 44*(8), 1344-1360.

Robinson, G. A., Spooner, D., & Harrison, W. J. (2015). Frontal dynamic aphasia in progressive supranuclear palsy: Distinguishing between generation and fluent sequencing of novel thoughts. *Neuropsychologia, 77*, 62-75.

Roccaforte, W. H., Burke, W. J., Bayer, B. L., & Wengel, S. P. (1992). Validation of a telephone version of the mini-mental state examination. *Journal of the American Geriatrics Society, 40*(7), 697-702.

Rönnlund, M., Nyberg, L., Bäckman, L., & Nilsson, L. (2005). Stability, growth, and decline in adult life span development of declarative memory: Cross-sectional and longitudinal data from a population-based study. *Psychology and Aging, 20*(1), 3-18.

Ruijs, T. (2018). *The Dutch Telephone-based Language-Battery: Evaluation of language in brain tumor patients after surgery* [Unpublished master's thesis]. University of Amsterdam.

Satoer, D., Kloet, A., Vincent, A., Dirven, C., & Visch-Brink, E. (2014). Dynamic aphasia following low-grade glioma surgery near the supplementary motor area: A selective spontaneous speech deficit. *Neurocase, 20*(6), 704-716.

Satoer, D., Vincent, A., Ruhaak, L., Smits, M., Dirven, C., & Visch-Brink, E. (2018). Spontaneous speech in patients with gliomas in eloquent areas: Evaluation until 1 year after surgery. *Clinical Neurology and Neurosurgery, 167*, 112-116.

Satoer, D., De Witte, E., Bulte, B., Bastiaanse, R., Smits, M., Vincent, A., Mariën, P., & Visch- Brink, E. (2019). *Diagnostic Instrument for Mild Aphasia (DIMA): Standardization and clinical application.* Unpublished manuscript. Erasmus MC University Medical Center, Rotterdam, The Netherlands.

Satoer, D., Piai, V., van Dijk, K., Ruijs, T., Visch-Brink, E., & De Witte, E. (2020a). *TeleTaalTest-NL: Telefonische test voor mensen met neurologische taalstoornissen.* Unpublished manuscript. Erasmus MC University Medical Center, Rotterdam, The Netherlands.

Satoer, D., Akinina, Y., de Kok, D., & Bastiaanse, R. (2020b). The Token Test App. Second edition: Dutch version. Groningen, The Netherlands: The University of Groningen.

Schwartz, B. L., & Metcalfe, J. (2011). Tip-of-the-tongue (TOT) states: Retrieval, behavior, and experience. *Memory & Cognition, 39*(5), 737-749.

Sierpowska, J., Gabarrós, A., Fernandez-Coello, A., Camins, A., Castañer, S., Juncadella, M., Moris, J., & Rodríguez-Fornells, A. (2017). Words are not enough: Nonword repetition as an indicator of arcuate fasciculus integrity during brain tumor resection. *Journal of Neurosurgery, 126*, 435-445.

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The Journal of the Acoustical Society of America, 26*(2), 212-215.

Talacchi, A., Santini, B., Casagrande, F. Alessandrini, F., Zoccatelli, G., & Squintani, G. M. (2013). Awake surgery between art and science. Part I: Clinical and operative settings. *Functional Neurology, 28*(3), 205-221.

Van den Berg, E., Ruis, C., Biessels, G. J., Kappelle, L. J., & van Zandvoort, J. E. (2012). The Telephone Interview for Cognitive Status (Modified): Relation with a comprehensive neuropsychological assessment. *Journal of Clinical and Experimental Neuropsychology, 34*(6), 598-605.

Van Dijk, K. (2018). *TeleLanguage-NL: Telefonisch gebaseerde taaltest batterij voor mensen met een hersentumor* [Unpublished master's thesis]. Radboud University Nijmegen.

Verhage, F. (1964). *Intelligentie en leeftijd: Onderzoek bij Nederlanders van 12 tot 77 jaar*. Assen, The Netherlands: Van Gorkum.

Visch-Brink, E. G., van de Sandt-Koenderman, M., & el Hachioui, H. (2010). *ScreeLing: Handleiding.* Houten, The Netherlands: Bohn Stafleu van Loghum.

Visch-Brink, E., & el Hachioui, H. (2013). *Afasie Bedside Check (ABC)*. Rotterdam, The Netherlands: Erasmus MC Rotterdam.

Visch-Brink, E., Vandenborre, D., de Smet, H. J., & Mariën, P. (2014). *Comprehensive Aphasia Test - Nederlandse bewerking - Handleiding.* Amsterdam, The Netherlands: Pearson.

Webb, A. J. S., Pendlebury, S. T., Li, L., Simoni, M., Lovett, N., Mehta, Z., & Rothwell, P. M. (2014). Validation of the Montreal Cognitive Assessment versus Mini-Mental State Examination against hypertension and hypertensive arteriopathy after Transient Ischemic Attack or minor stroke. *Stroke, 45*(11), 3337-3342.

Wilson, S. M., Eriksson, D. K., Schneck, S. M., & Lucanie, J. M. (2018). A quick aphasia battery for efficient, reliable, and multidimensional assessment of language function. *PloS One, 13*(2):e0192773.

Yochim, B. P., Beaudreau, S. A., Kaci Fairchild, J., Yutsis, M. V., Raymond, N., Friedman, L., & Yesavage, J. (2015). Verbal naming test for use with older adults: Development and initial validation. *Journal of the International Neuropsychology Society, 21*(3), 239-248.

Zietemann, V., Kopczak, A., Müller, C., Wollenweber, F. A., & Dichgans, M. (2017). Validation of the Telephone Interview of Cognitive Status and Telephone Montreal Cognitive Assessment against detailed cognitive testing and clinical diagnosis of mild cognitive impairment after stroke. *Stroke, 48*(11), 2952-2957.

Appendix A: Healthy control group DIMA

Table A1. Demographic characteristics of the DIMA healthy control group.

Demographics	Mean	SD	Range
Age in years	50.20	19.76	18.07-85.37
Education in years	13.68	3.48	6-30
Demographics	Groups	Number of participants	Percentage
Gender	Female	105	49.1
	Male	109	50.9
Age group	18-54 years	116	54.2
	55-85 years	96	44.9
	Missing	2	0.9
Education level	≤ 12 years	98	45.8
	> 12 years	115	53.7
	Missing	1	0.5
Nationality	Belgian	109	50.9
	Dutch	105	49.1
Handedness	Left	18	8.4
	Right	183	85.5
	Ambidexter	9	4.2
	Missing	4	1.9

Table A2. Scores DIMA healthy control group.

Subtest (total score)	Mean	SD	Range
Phonological repetition (40)	37.38	1.48	16 - 40
• PHON A - three-syllabic words (10)	9.74	0.74	4 - 10
• PHON B - compounds (10)	9.66	0.91	4 - 10
• PHON C - nonwords (10)	9.01	1.47	2 - 10
• PHON D - sentences (10)	8.95	1.48	0 - 10
SEM A - Semantic odd picture out nouns (5)	4.49	0.90	0 - 5
SYN A - Sentence completion (10)	9.37	1.06	4 - 10
DIMA total score (only SEM A) (55)	51.20	4.82	24 - 55

Table A3. DIMA healthy control scores according to age and education in years.

Subtest (total score)	Mean				SD				Range			
	Age		Education in years		Age		Education in years		Age		Education in years	
	< 55	≥ 55	≤ 12	> 12	< 55	≥ 55	≤ 12	> 12	< 55	≥ 55	≤ 12	> 12
Phonological repetition (40)	38.56	35.98	36.57	38.07	1.80	4.77	4.08	3.19	32 - 40	16 - 40	16 - 40	19 - 40
• PHON A - three-syllabic words (10)	9.88	9.56	9.63	9.83	0.40	0.98	0.90	0.55	8 - 10	4 - 10	4 - 10	6 - 10
• PHON B - compounds (10)	9.86	9.42	9.51	9.79	0.42	1.23	1.14	0.62	8 - 10	4 - 10	4 - 10	7 - 10
• PHON C - nonwords (10)	9.43	8.50	8.67	9.29	0.90	1.83	1.70	1.18	6 - 10	2 - 10	2 - 10	3 - 10
• PHON D - sentences (10)	9.37	8.45	8.67	9.18	0.95	1.83	1.60	1.34	5 - 10	0 - 10	0 - 10	1 - 10
SEM A - Semantic odd picture out nouns (5)	4.78	4.11	4.37	4.58	0.47	1.14	0.99	0.82	3 - 5	0 - 5	0 - 5	0 - 5
SYN A - Sentence completion (10)	9.66	9.02	9.18	9.52	0.69	1.31	1.21	0.88	7 - 10	4 - 10	4 - 10	5 - 10
DIMA total score (only SEM A) (55)	52.99	49.07	50.04	52.19	2.20	6.09	5.44	3.97	44 - 55	24 - 55	24 - 55	31 - 55

Appendix B: Information letter TeleTaalTest-NL

FETC-GW-referentienummer: 6656609-01-02-2020

Informatie over deelname aan

Normeringsonderzoek: Telefonische taaltest voor mensen met een neurologische taalstoornis

1. Inleiding

Beste meneer/mevrouw,

Wij vragen u vriendelijk om mee te doen aan een wetenschappelijk onderzoek van de Universiteit Utrecht in samenwerking met het Erasmus MC - Universitair Medisch Centrum Rotterdam. Dit onderzoek is telefonisch en vindt niet plaats op locatie. U beslist zelf of u wilt meedoen. Voordat u de beslissing neemt, is het belangrijk om meer te weten over het onderzoek. Lees deze informatiebrief rustig door. Heeft u na het lezen van de informatie nog vragen? Dan kunt u terecht bij de onderzoeker. Op bladzijde 4 vindt u de contactgegevens. Indien u toch besluit om niet deel te nemen aan deze studie, dan hoeft u niets te doen. Dit onderzoek is getoetst door de Facultaire Ethische ToetsingsCommissie – Geesteswetenschappen (FETC-GW) van Universiteit Utrecht, te zien aan het FETC-GW referentienummer in de rechterbovenhoek.

2. Wat is de achtergrond en het doel van het onderzoek?

Patiënten die hersenletsel hebben opgelopen, zoals een beroerte, een traumatisch hersenletsel of een hersentumor(operatie), kunnen als gevolg daarvan last hebben van taalproblemen (afasie). De ernst van de taalstoornis kan variëren van milde problemen tot ernstige problemen. Om de aard en de ernst van de afasie vast te stellen worden mensen met verdenking van afasie ten gevolge van een neurologische stoornis op de polikliniek van bijvoorbeeld een ziekenhuis of een revalidatiecentrum gezien door een klinisch linguïst en/of logopedist die een neurolinguïstisch test-onderzoek afneemt. Op basis hiervan kan de afasie goed in kaart worden gebracht en/of kan gekeken worden of er sprake is van vooruitgang na taaltherapie. Gewoonlijk zijn “face-to-face” intake- en vervolgafspraken door de relatief korte reisafstanden in Nederland geen probleem. Echter, door onvoorziene situaties zoals COVID-19, wanneer alle niet-spoedeisende afspraken worden afgezegd, worden ook deze afspraken geannuleerd. Om de poliklinische afspraken op een alternatieve manier toch te kunnen voortzetten zodat er toch gescreend kan worden op afasie en het zorgproces van de patiënt gewaarborgd kan worden, is de TeleTaalTest-NL ontwikkeld. De test is gebaseerd op de Amerikaanse TeleLanguage Test, een taaltest die oorspronkelijk gemaakt is voor hersentumor patiënten die om verschillende redenen geen toegang hebben tot een ziekenhuis en/of videobellen. Echter, deze telefonische taaltest zou ook relevant kunnen zijn voor mensen met taalstoornissen ten gevolge van een andere neurologische aandoening, zoals een beroerte. Een eerdere versie van de TeleTaalTest-NL is al bij gezonde mensen afgenomen om als

referentiegroep te dienen voor patiënten met een neurologische aandoening. Een studie met gegevens van gezonde mensen (normeringsonderzoek) is van belang om een betrouwbare vergelijking te kunnen maken om te zien of er sprake is van een taalprobleem bij mensen die wel een hersenaandoening hebben. Aangezien deze nieuwe versie van de test verbeteringen bevat en het aantal geteste gezonde mensen relatief laag was, willen wij met deze studie meer gegevens verzamelen van mensen zonder hersenaandoening.

3. Hoe wordt het onderzoek uitgevoerd?

De TeleTaalTest-NL zal telefonisch worden afgenomen. U wordt verzocht om in een stille ruimte zonder te veel afleidingen te gaan zitten. Voorafgaand aan het onderzoek wordt er een korte vragenlijst afgenomen met betrekking tot uw algemene gegevens. Deze informatie is nodig om eventuele variatie in testresultaten te kunnen verklaren. Alleen tijdens de testen zullen audio-opnames gemaakt worden zodat de scoring van uw testresultaten nauwkeurig kan worden nagelopen door de onderzoekers. De echte test bestaat uit één korte screening en vier korte onderdelen die elk een ander aspect van de taalvaardigheid testen. De door u behaalde scores zullen met de scores van de andere gezonde deelnemers vergeleken worden om vast te stellen of het huidige testmateriaal nog verbetering nodig heeft. Alle scores zullen uiteindelijk ook vergeleken worden met de eerder verzamelde scores van bestaande afasietesten om vast te stellen of de TeleTaalTest-NL dezelfde taalvaardigheden meet en dus een goed alternatief is. Indien u wilt deelnemen aan het onderzoek vragen wij u eerst om de toestemmingsverklaring ondertekend via e-mail terug te sturen. Na ontvangst van dit formulier kan een afspraak gemaakt worden om de taaltest telefonisch bij u af te nemen.

4. Wat wordt er van u verwacht?

Als u zou willen deelnemen, zal een medewerker van het onderzoeksteam éénmalig telefonisch een taaltest afnemen. We verwachten dat het afnemen van deze test 15-20 minuten duurt.

5. Wat zijn mogelijke voor- en nadelen van deelname aan dit onderzoek?

U heeft zelf geen voordeel van deelname aan dit onderzoek. Voor de toekomst kan het onderzoek wel nuttige gegevens opleveren. Zo draagt het onderzoek bij aan het ontwikkelen van testmateriaal voor mensen met een neurologische taalstoornis.

Mogelijke nadelen zijn:

- Het beantwoorden van de vragen en het uitvoeren van de test kost u tijd.
- De tests kunnen moeilijker zijn dan verwacht, wat confronterend kan zijn.

6. Vrijwillige deelname

Deelname is vrijwillig. Als u toch besluit niet deel te nemen, dan hoeft u verder niets te doen. U hoeft niets te tekenen en u bent ook niet verplicht om te zeggen waarom u niet wilt deelnemen. Als u wel deelneemt, kunt u zich altijd bedenken en stoppen op ieder gewenst

moment – ook tijdens het onderzoek. Het onderzoek kan ook door de onderzoeker worden beëindigd wanneer hier een geldige reden voor is. Bovendien kunt u nadat u heeft meegedaan nog uw toestemming intrekken. Indien u daarvoor kiest, zullen uw onderzoeksgegevens niet meegenomen worden in de analyses. Uw onderzoeksgegevens kunnen echter niet meer verwijderd worden als deze al zijn geanalyseerd of als uw onderzoeksgegevens niet meer tot u te herleiden zijn, zie hieronder.

7. Wat gebeurt er met de verzamelde gegevens?
De gegevens die in het kader van dit onderzoek over u verzameld worden, zullen vertrouwelijk worden behandeld. Uw persoonsgegevens (naam, adres en andere privacygevoelige gegevens) worden beheerd door één persoon, Naomi Legius. Mocht u uw persoonsgegevens willen corrigeren of laten verwijderen, dan kunt u dit doen door contact op te nemen via: naomilegius@gmail.com. Uw persoonsgegevens zullen niet verstrekt worden aan anderen dan de onderzoekers die direct bij dit project betrokken zijn.

Tijdens uw deelname aan het onderzoek worden gegevens verzameld. Deze gegevens krijgen een code die alleen met een sleutel naar u te herleiden is. De onderzoeksgegevens zijn dus gecodeerd en losgekoppeld van uw persoonlijke gegevens. In eventuele wetenschappelijke publicaties zal uw naam dan ook op geen enkele manier zijn terug te vinden. Sommige mensen mogen uw onderzoeksgegevens inzien. Dit is om te controleren of het onderzoek goed en betrouwbaar is. Zij zien alleen de code, nooit uw naam.

De personen die inzage kunnen krijgen in uw onderzoeksgegevens zijn:
- De medewerkers van het onderzoeksteam

Wij zijn verplicht de gecodeerde onderzoeksgegevens minimaal 10 jaar te bewaren. Dit is nodig om het onderzoek te kunnen controleren en daarvoor geeft u toestemming als u meedoet aan dit onderzoek. Als u dat niet wilt, kunt u niet meedoen aan dit onderzoek. Uw gegevens worden opgeslagen en bewaard op een door de Universiteit Utrecht beveiligde server. Het zou kunnen dat uw gegevens ook nuttig zijn voor toekomstige onderzoeken. De gecodeerde onderzoeksgegevens en het geluidsmateriaal (dit is niet anoniem) kunnen dan worden gedeeld met en/of worden hergebruikt door andere wetenschappers om eventueel andere onderzoeksvragen mee te beantwoorden. Op de toestemmingsverklaring kunt u apart aangegeven of u toestemming geeft voor het gebruiken van uw gegevens en geluidsmateriaal in toekomstige onderzoeken. Vanzelfsprekend blijft de vertrouwelijkheid van de gegevens altijd gelden. De sleutel waarmee gecodeerde gegevens naar u persoonlijk te herleiden zijn, wordt zo spoedig mogelijk vernietigd.

8. Is er een vergoeding wanneer u besluit aan dit onderzoek mee te doen?
Er is geen vergoeding verbonden aan deelname aan dit onderzoek.

9. Goedkeuring van dit onderzoek

De Facultaire Ethische ToetsingsCommissie – Geesteswetenschappen (FETC-GW) heeft dit onderzoek goedgekeurd. Wanneer u een klacht wilt indienen over de procedure omtrent dit onderzoek, dan kunt u contact opnemen met de secretaris van de FETC-GW, e-mail: fetc-gw@uu.nl, of met de functionaris voor gegevensbescherming van de Universiteit Utrecht, e-mail: privacy@uu.nl.

10. Meer informatie over dit onderzoek?

Indien u na zorgvuldige overweging besluit deel te nemen aan dit wetenschappelijk onderzoek, dan vragen we u om de toestemmingsverklaring digitaal te ondertekenen (door middel van een digitale afbeelding van uw handtekening), te dateren en het formulier vanaf een persoonlijk e-mailadres retour te sturen. Een ingescande getekende verklaring is ook akkoord. Als u meer informatie over het onderzoek wilt, zijn wij natuurlijk altijd bereid u te woord te staan. U kunt Dr. Satoer, de coördinator van het onderzoek, op werkdagen bereiken per e-mail (zie contactgegevens onderaan de brief).

Met vriendelijke groet,

Dr. D.D. Satoer,
Onderzoeksleider
Afdeling Neurochirurgie
Erasmus MC Rotterdam
Postbus 2040
3000 CA Rotterdam
E-mail: d.satoer@erasmusmc.nl
Telefoon: 06-50008179

Mede namens,

Prof. dr. S. Avrutin
Hoogleraar comparative psycholinguistics
Afdeling Geesteswetenschappen
Universiteit Utrecht
Trans 10
3512 JK Utrecht
E-mail: S.Avrutin@uu.nl
Telefoon: 06-52373849

N.A.A.M. Legius
Student onderzoeksmaster taalkunde
Universiteit Utrecht
E-mail: naomilegius@gmail.com
Telefoon: 06-15621193

11. Bijlagen:

Bijlage 1: Toestemmingsverklaring.

Appendix C: Informed consent form TeleTaalTest-NL

FETC-GW-referentienummer: 6656609-01-02-2020

TOESTEMMINGSVERKLARING voor deelname aan:

Normeringsonderzoek: Telefonische taaltest voor mensen met neurologische taalstoornissen

Ik bevestig:

(1) dat ik via de informatiebrief naar tevredenheid over het onderzoek ben ingelicht;
(2) dat ik in de gelegenheid ben gesteld om vragen over het onderzoek te stellen en dat mijn eventuele vragen naar tevredenheid zijn beantwoord;
(3) dat ik gelegenheid heb gehad om grondig over deelname aan het onderzoek na te denken;
(4) dat ik uit vrije wil deelneem.

Ik stem er mee in dat:

(5) de verzamelde gegevens voor wetenschappelijke doelen worden verkregen en bewaard zoals in de informatiebrief vermeld staat;
(6) de verzamelde, gecodeerde, onderzoeksgegevens door wetenschappers kunnen worden gedeeld en/of worden hergebruikt om eventueel andere onderzoeksvragen mee te beantwoorden;
[ ] Ja, daar ga ik mee akkoord. **[ ]** Nee, niet akkoord.

(7) er voor wetenschappelijke doeleinden geluidsopnamen worden gemaakt (zie voor delen en hergebruik van deze gegevens onderstaande box).

Ik begrijp dat:

- ik het recht heb om mijn toestemming voor het gebruik van data in te trekken, zoals vermeld staat in de informatiebrief.

Naam deelnemer: _____ Geboortedatum: ___ / ___ / ____ (dd/mm/jjjj)

Handtekening: _____ Datum, plaats: ___ / ___ / ____, _____

Verklaring omtrent hergebruik van gegevens: **in te vullen nadat de dataverzameling heeft plaatsgevonden** (s.v.p. aankruisen wat van toepassing is, en onderaan ondertekenen).

- Gaat u er mee akkoord dat het verzamelde audio materiaal, dit is dus niet anoniem, gedeeld wordt met andere onderzoekers voor onderzoeksdoeleinden?

  [ ] Ja, daar ga ik mee akkoord. **[ ]** Nee, niet akkoord.

- Soms worden geluidsopnames ook getoond in wetenschappelijke lezingen of lessen. Gaat u ermee akkoord dat het gemaakte audio materiaal gebruikt wordt voor deze doeleinden?

  [ ] Ja, daar ga ik mee akkoord. **[ ]** Nee, niet akkoord.

Handtekening: _____

In te vullen door de uitvoerend onderzoeker:

Naam: _____

Ik verklaar dat ik bovengenoemde deelnemer heb uitgelegd wat deelname aan het onderzoek inhoudt.

Datum: ___ / ___ / _____ (dd/mm/jjjj)

Handtekening:

Appendix D: Score form Edinburgh Handedness Inventory (EHI)

Edinburgh Handedness Inventory

Oldfield (1971)

Wilt u uw initialen invullen:

Uw geboortedatum:

De datum van vandaag:

Instructie:
Duid a.u.b. uw handvoorkeur aan voor de volgende activiteiten door een + te zetten in de juiste kolom. Als uw voorkeur zo sterk is dat u nooit zou proberen de andere hand te gebruiken, tenzij u er absoluut toe verplicht wordt zet dan ++. Als het u werkelijk om het even is, zet dan een + in beide kolommen. Bij sommige van de activiteiten heeft u beide handen nodig. In deze gevallen wordt de handvoorkeur gevraagd voor het voorwerp of activiteit tussen haakjes. Probeer a.u.b. alle vragen te beantwoorden, en laat het enkel blanco wanneer u geen ervaring hebt met de taak.

		L	R
1	Schrijven		
2	Tekenen		
3	Werpen		
4	Schaar		
5	Tandenborstel		
6	Mes (zonder vork)		
7	Lepel		
8	Borstel (bovenste hand)		
9	Lucifer aanstrijken (lucifer)		
10	Jampot openen (deksel)		
	Totaal	L=	R=

Percentage: $\frac{R-L}{R+L}$ x 100 van _____

	Interpretatie
Linkshandig	$R < -40$
Ambidexter	$-40 \leq R \leq +40$
Rechtshandig	$R > +40$

Scoreformulier TeleTaalTest-NL

Algemene gegevens:

Deelnemer:

Geboortedatum / leeftijd:

Geslacht:

Datum:

Hoogste opleidingsniveau (in jaren):

(Voormalig) beroep:

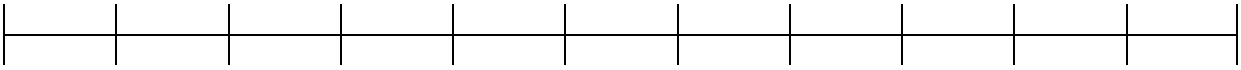
Gehoor: goed / matig / slecht

Gehoorapparaat: ja / nee

Aan welk oor heb je de telefoon?

In te vullen na afloop van de testafname:

Geluidskwaliteit telefoongesprek:



Uitleg vooraf:
Deze taaltest bestaat uit verschillende onderdelen. Voor elke test wordt er een korte uitleg gegeven. Heeft u tussendoor nog vragen of opmerkingen, dan kunt u dat gewoon zeggen. U kunt op elk moment stoppen als u dat wilt.

Screening begrip

Instructie: We beginnen met 5 ja/nee vragen. Ik stel u een vraag waarop u met ‘ja’ of met ‘nee’ kunt antwoorden. Als u het antwoord niet weet, zegt u ‘weet ik niet’.

Scoring: 1 punt bij een goed antwoord, 0 bij een fout antwoord. Alle vragen moeten goed beantwoord worden om door te kunnen met de overige sub-testen.

	Vraag	Antwoord	Score		Opmerking
1.	Is uw achternaam ‘Jansen’?	Nee	0	1	
2.	Bent u een vrouw?*	Ja	0	1	
opt2	Bent u een kind?	Nee	0	1	
opt3	Bent u 18 jaar of ouder?	Ja	0	1	
3.	Is het nu winter?*	Nee	0	1	
4.	Komt de maand *mei* vóór de maand *juni?*	Ja	0	1	
5.	Is bloemkool een fruitsoort?	Nee	0	1	
* afstemmen op het geslacht van de deelnemer en tegenovergestelde seizoen vragen					

Woordvinding

Instructie: hoort een zin. Zegt u het woord dat daarbij hoort. Let goed op de betekenis. Als u het antwoord niet weet, zeg dan 'weet ik niet'. We gaan eerst een keer oefenen.

Scoring: Als de deelnemer niet op het woord kan komen, worden er 0 punten genoteerd; fonologische of semantische cues worden niet gegeven**.** Als het woord in een zin wordt gezegd of in een zinsdeel wordt het ook goed gerekend. Bijvoorbeeld bij 'bed': 'ik slaap in een bed' of 'in een bed'.

Punten	**Antwoord**
2 punten	Juiste benoeming binnen 4 seconden*
1 punt	Juiste benoeming na 4 seconden
0 punten	Foutief benoeming/ 'weet ik niet'
Afbreken na 5 opeenvolgende foutieve antwoorden	

Doelwoord		**Frequentie***	**Vraag**	**Score**	**Opmerking**
Oefen item	*Bed*	239.93	*Waarin slaap je 's nachts?*		
1.	Vader	795.62	De man heeft een dochter, hij is dus een…	0 1 2	
2.	Uur	348.69	Er zitten 60 minuten in een...	0 1 2	
3.	Hoofd	274.05	Je oren zitten aan je…?	0 1 2	
4.	School	246.95	Waar krijgen kinderen les?	0 1 2	
5.	Water	244.50	Wat komt er uit de kraan?	0 1 2	
6.	Leger	107.98	Hoe noem je meerdere soldaten bij elkaar?	0 1 2	
7.	Dansen	102.91	Hoe heet bewegen op muziek ook wel?	0 1 2	
8.	Verliezen	96.78	Bij een wedstrijd wint het ene team, wat doet het andere team?	0 1 2	
9.	Lachen	88.89	Wat doe je als iets heel grappig is?	0 1 2	
10.	Kussen	45.62	Waar ligt je hoofd op als je slaapt?	0 1 2	
11.	Films	40.86	Wat draaien ze in de bioscoop?	0 1 2	
12.	Melk	39.70	Welke drank komt er van de koe?	0 1 2	
13.	Piloot	30.12	Hoe noem je iemand die een vliegtuig bestuurt?	0 1 2	
14.	Kasteel	27.60	Waarin wonen een prins en prinses?	0 1 2	

16.	Hamer	8.55	Waarmee sla je een spijker in de muur?	0 1 2	
18.	Bakkerij	2.47	Waar worden broden en taarten gemaakt?	0 1 2	
19.	IJsbaan	0.94	Hoe noem je de plek waar schaatsers trainen?	0 1 2	
II.3	Zoeken	273.23	Wanneer je verstoppertje speelt ga je iemand anders…?	0 1 2	
II.4	Spelen	247.77	Wat doet een kind met poppen?	0 1 2	
II.12	Zwemmen	39.47	Wat doen mensen graag bij warm weer in de zee?	0 1 2	
II.14	Tellen	27.35	Wat ben je aan het doen als je getallen op volgorde opnoemt?	0 1 2	
II.19	Hypnotiseren	0.94	Wat doet iemand die mensen in trance brengt?	0 1 2	
II.20	Snoeien	0.75	Hoe noem je het wanneer een tuinman takken van een boom knipt?	0 1 2	
III.3	Vergeten	230.44	Hoe noem je dat als je iets even niet meer weet?	0 1 2	
R.1	Geld	793.61	Contant betalen doe je met…	0 1 2	
R.2	Politie	346.17	Een agent is onderdeel van de…	0 1 2	
R.3	Bellen	228.77	Hoe noem je dat als je elkaar spreekt via de telefoon?	0 1 2	
R.4	Muziek	107.46	Wat hoor je op de radio?	0 1 2	
R.5	Liegen	60.33	Hoe heet het als je de waarheid niet vertelt?	0 1 2	
R.6	Voeren/ voederen (BE)	36.29	Hoe noem je het wanneer eenden stukjes brood krijgen?	0 1 2	
R.6V	Voeren/ voederen (BE)	36.29	Hoe noem je het wanneer iemand eenden stukjes brood geeft?	0 1 2	
R.7	Waarschuwen	27.42	Hoe noem je het als je zegt dat iemand moet opletten?	0 1 2	
R.7V	Waarschuwen	27.42	Hoe noem je het als je tegen iemand zegt dat hij moet opletten?	0 1 2	

* Per miljoen woorden

Fonologie: Nazeggen

Instructie: Ik zeg een paar woorden en zinnen. Zegt u mij maar na. Er zijn ook een paar onzinwoorden bij. We gaan beginnen. We zullen eerst een keer oefenen (kondig de niet-bestaande woorden en zinnen tijdens de test aan)

Scoring:

Score	**Punten**	**Antwoord**
Woorden	1	Correcte herhaling
	0	Foutieve herhaling
Zinnen	De score is gelijk aan het aantal inhoudswoorden in de zin. *Mogelijke fonematische, dysarthrische en spraak-apractische antwoorden worden niet als foutief gescoord aangezien dat via de telefoon lastig te beoordelen is.*	

	Bestaande woorden	**Score**	**Opmerking**
Oefenitem	*Legende*		
1.	Radio	0 1	
2.	Kanarie	0 1	
3.	Trommelvlies	0 1	
4.	Kleuterschool	0 1	
	Niet-bestaande woorden		
5.	Ánamo	0 1	
6.	Frimótika	0 1	
	Doelzin		
Oefen item	*Hij krijgt een geschenk*		
7.	De jongens bouwen een brug (3)	0 1 2 3	
8.	De kinderen fietsen naar huis (3)	0 1 2 3	
9.	De vrouw vergat de vissen te voeren (4)	0 1 2 3 4	
10.	Waarom winnen wij nooit een wedstrijd? (2)	0 1 2	

Semantiek: Semantische selectietest

Instructie: U krijgt straks 3 woorden te horen. Let op de betekenis. Eén woord hoort niet in het rijtje thuis. Welk woord is dat? En waarom past het niet bij de andere woorden? Als u het niet weet, zegt u ‘weet ik niet’. We zullen eerst een keer oefenen.

Scoring:

Punten	**Antwoord**
2 punten	Juiste antwoord met juiste uitleg
1 punt	Juiste antwoord zonder uitleg
0 punten	Foutief antwoord / ‘weet ik niet’

	Woord 1	**Woord 2**	**Woord 3**	**Doelwoord**	**Score**	**Opmerking**
Oefenitem	*Vlieg*	*Schip*	*Kever*	*Schip*		
1.	Boek	Zee	Krant	Zee	0 1 2	
2.	Sla	Wortel	Huis	Huis	0 1 2	
3.	Kerk	Hond	Paard	Kerk	0 1 2	
4.	Piano	Fluit	Koe	Koe	0 1 2	
5.	Zomer	Schaar	Winter	Schaar	0 1 2	
6.	Peer	Appel	Land	Land	0 1 2	
7.	IJs	Kachel	Oven	IJs	0 1 2	
8.	Laars	Schoen	Buik	Buik	0 1 2	
9.	Emmer	Brief	Gieter	Brief	0 1 2	
10.	Maan	Vork	Lepel	Maan	0 1 2	

Syntaxis: Zinnen afmaken

Instructie: Ik zeg het begin van een zin. Zegt u mij maar na en maak de zin af. We gaan eerst een keer oefenen.

Scoring:

Punten	**Antwoord**
2 punten	Een vloeiende en syntactisch correcte zin
1 punt	Een zin met een duidelijke hapering, twijfel of herhaling
0 punten	Een syntactisch incorrecte zin/ ‘weet ik niet’

	Doel zin	**Score**	**Opmerking**
Oefen item	*Ik was mijn handen met…*		
	Iedere maand …		
1.	Hij valt van...	0 1 2	
2.	De pan staat op...	0 1 2	
3.	Ik snij met...	0 1 2	
4.	De vrouw probeerde...	0 1 2	
5.	De jongen gelooft dat...	0 1 2	
6.	De man in de winkel...	0 1 2	
7.	Elke week...	0 1 2	
8.	Zij belde niet omdat...	0 1 2	
9.	Toen de vaas brak...	0 1 2	
10.	Het meisje benadrukt dat…	0 1 2	

Scoreformulier TeleTaalTest-NL

Algemene gegevens:

Deelnemer:

Geboortedatum / leeftijd:

Geslacht:

Datum:

Hoogste opleidingsniveau (in jaren):

(Voormalig) beroep:

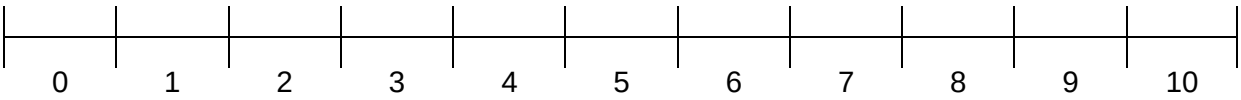
Gehoor: goed / matig / slecht

Gehoorapparaat: ja / nee

Aan welk oor heeft u de telefoon?

In te vullen na afloop van de testafname:

Geluidskwaliteit telefoongesprek:



Uitleg vooraf:

Deze taaltest bestaat uit verschillende onderdelen. Voor elke test wordt er een korte uitleg gegeven. Heeft u tussendoor nog vragen of opmerkingen, dan kunt u dat gewoon zeggen. U kunt op elk moment stoppen als u dat wilt.

Screening begrip

Instructie: We beginnen met 5 ja/nee vragen. Ik stel u een vraag waarop u met ‘ja’ of met ‘nee’ kunt antwoorden. Als u het antwoord niet weet, zegt u ‘weet ik niet’.

Scoring: 1 punt bij een goed antwoord, 0 bij een fout antwoord. Alle vragen moeten goed beantwoord worden om door te kunnen met de overige sub-testen. Wanneer de vraag niet goed verstaan is, dient de testleider de volledige vraag te herhalen en het aantal herhalingen bij te houden in de daarvoor bestemde kolom.

	Vraag	Antwoord	Score		Opmerking	Herhaling
1.	Is uw achternaam ‘Jansen’?	Nee	0	1		
2.	Bent u kind?	Nee	0	1		
3.	Is het nu winter?*	Nee	0	1		
4.	Komt de maand *mei* vóór de maand *juni?*	Ja	0	1		
5.	Is bloemkool een fruitsoort?	Nee	0	1		
* het tegenovergestelde seizoen vragen						

Woordvinding

Instructie: U hoort een zin. Zegt u het woord dat daarbij hoort. Let goed op de betekenis. Als u het antwoord niet weet, zeg dan ‘weet ik niet’. We gaan eerst een keer oefenen.

Scoring: Als de deelnemer niet op het woord kan komen, worden er 0 punten genoteerd; fonologische of semantische cues worden niet gegeven. Als het woord in een zin wordt gezegd of in een zinsdeel wordt het ook goed gerekend. Bijvoorbeeld bij ‘bed’: ‘ik slaap in een bed’ of ‘in een bed’. In alle gevallen telt het laatste antwoord van de deelnemer. Wanneer de vraag niet goed verstaan is, dient de testleider de volledige vraag te herhalen en het aantal herhalingen bij te houden in de daarvoor bestemde kolom.

Punten	**Antwoord**
2 punten	Juiste benoeming binnen 4 seconden*
1 punt	Juiste benoeming na 4 seconden
0 punten	Foutief benoeming/ ‘weet ik niet’
Afbreken na 5 opeenvolgende foutieve antwoorden	

Test items

	Vraag	**Doelwoord**	**Frequentie***	**Score**	**Opmerking**	**Herhaling**
Oefen item	*Waarin slaap je ’s nachts?*	*Bed*	239.93	0 1 2		
1.	Er zitten 60 minuten in een…	Uur	348.69	0 1 2		
2.	Wat komt er uit de kraan?	Water	244.50	0 1 2		
3.	Hoe noem je dat als je elkaar spreekt via de telefoon?	Bellen	228.77	0 1 2		
4.	Wat hoor je op de radio?	Muziek	107.46	0 1 2		
5.	Bij een wedstrijd wint het ene team, wat doet het andere team?	Verliezen	96.78	0 1 2		
6.	Wat doe je als iets heel grappig is?	Lachen	88.89	0 1 2		
7.	Wat doen mensen graag bij warm weer in de zee?	Zwemmen	39.47	0 1 2		
8.	Hoe noem je iemand die een vliegtuig bestuurt?	Piloot	30.12	0 1 2		
9.	Waar worden broden en taarten gemaakt?	Bakkerij	2.47	0 1 2		
10.	Hoe noem je het wanneer een tuinman takken van een boom knipt?	Snoeien	0.75	0 1 2		

* Per miljoen woorden

Bij twijfel over de eerste afname kan er een tweede set van tien items worden afgenomen.

Reserve items

	Vraag	**Doelwoord**	**Frequentie***	**Score**	**Opmerking**	**Herhaling**
1.	Een agent is onderdeel van de…	Politie	346.17	0 1 2		
2.	Waar krijgen leerlingen les?	School	246.95	0 1 2		
3.	Wat doet een kind met poppen?	Spelen	247.77	0 1 2		
4.	Hoe noem je meerdere soldaten bij elkaar?	Leger	107.98	0 1 2		
5.	Hoe heet bewegen op muziek ook wel?	Dansen	102.91	0 1 2		
6.	Hoe heet het als je de waarheid niet vertelt?	Liegen	60.33	0 1 2		
7.	Hoe noem je het wanneer iemand eenden stukjes brood geeft?	Voeren/ voederen (BE)	36.29	0 1 2		
8.	Waarin wonen een prins en prinses?	Kasteel	27.60	0 1 2		
9.	Waarmee sla je een spijker in de muur?	Hamer	8.55	0 1 2		
10.	Wat doet iemand die mensen in trance brengt?	Hypnotiseren	0.94	0 1 2		

* Per miljoen woorden

Fonologie: Nazeggen

Instructie: Ik zeg een paar woorden en zinnen. Zegt u mij maar na. Er zijn ook een paar onzinwoorden bij. We gaan beginnen. We zullen eerst een keer oefenen (kondig de niet-bestaande woorden en zinnen tijdens de test aan)

Scoring: Wanneer het woord of de zin niet goed verstaan is, dient de testleider het volledige item te herhalen en het aantal herhalingen bij te houden in de daarvoor bestemde kolom.

Score	**Punten**	**Antwoord**
Woorden	1	Correcte herhaling
	0	Foutieve herhaling
Zinnen	De score is gelijk aan het aantal inhoudswoorden in de zin. *Mogelijke fonematische, dysarthrische en spraak-apractische antwoorden worden niet als foutief gescoord aangezien dat via de telefoon lastig te beoordelen is.*	

	Bestaande woorden	**Score**	**Opmerking**	**Herhaling**
Oefenitem	*Legende*			
1.	Radio	0 1		
2.	Kanarie	0 1		
3.	Trommelvlies	0 1		
4.	Kleuterschool	0 1		
	Niet-bestaande woorden			
5.	Ánamo	0 1		
6.	Prekamlédi	0 1		
	Doelzin			
Oefenitem	*Hij krijgt een geschenk*			
7.	De jongens bouwen een brug (3)	0 1 2 3		
8.	De kinderen fietsen naar huis (3)	0 1 2 3		
9.	De vrouw vergat de vissen te voeren (4)	0 1 2 3 4		
10.	Waarom winnen wij nooit een wedstrijd? (2)	0 1 2		

Semantiek: Semantische selectietest

Instructie: U krijgt straks 3 woorden te horen. Let op de betekenis. Eén woord hoort niet in het rijtje thuis. Welk woord is dat? En waarom past het niet bij de andere woorden? Als u het niet weet, zegt u ‘weet ik niet’. We zullen eerst een keer oefenen.

Scoring: Wanneer de reeks of een deel daarvan niet goed verstaan is, dient de testleider het volledige item te herhalen en het aantal herhalingen bij te houden in de daarvoor bestemde kolom.

Punten	**Antwoord**
2 punten	Juiste antwoord met juiste uitleg
1 punt	Juiste antwoord zonder uitleg
0 punten	Foutief antwoord / ‘weet ik niet’

	Woord 1	**Woord 2**	**Woord 3**	**Doelwoord**	**Score**	**Opmerking**	**Herhaling**
Oefenitem	*Vlieg*	*Schip*	*Kever*	*Schip*			
1.	Boek	Zee	Krant	Zee	0 1 2		
2.	Sla	Wortel	Huis	Huis	0 1 2		
3.	Kerk	Hond	Paard	Kerk	0 1 2		
4.	Piano	Fluit	Koe	Koe	0 1 2		
5.	Zomer	Schaar	Winter	Schaar	0 1 2		
6.	Peer	Appel	Land	Land	0 1 2		
7.	IJs	Kachel	Oven	IJs	0 1 2		
8.	Laars	Schoen	Buik	Buik	0 1 2		
9.	Emmer	Brief	Gieter	Brief	0 1 2		
10.	Maan	Vork	Lepel	Maan	0 1 2		

Syntaxis: Zinnen afmaken

Instructie: Ik zeg het begin van een zin. Zegt u mij maar na en maak de zin af. We gaan eerst een keer oefenen.

Scoring: Wanneer de zin niet goed verstaan is, dient de testleider de volledige zin te herhalen en het aantal herhalingen bij te houden in de daarvoor bestemde kolom.

Punten	**Antwoord**
2 punten	Een vloeiende en syntactisch correcte zin
1 punt	Een zin met een duidelijke hapering, twijfel (een pauze langer dan 4 seconden) of herhaling
0 punten	Een syntactisch incorrecte zin/ 'weet ik niet'

	Doel zin	**Score**	**Opmerking**	**Herhaling**
Oefen item	*Ik was mijn handen met…*			
Oefen item	*Iedere maand …*			
1.	Hij valt van...	0 1 2		
2.	De pan staat op...	0 1 2		
3.	Ik snij met...	0 1 2		
4.	De vrouw probeerde...	0 1 2		
5.	De jongen gelooft dat...	0 1 2		
6.	De man in de winkel...	0 1 2		
7.	Elke week...	0 1 2		
8.	Zij belde niet omdat...	0 1 2		
9.	Toen de vaas brak...	0 1 2		
10.	Het meisje benadrukt dat…	0 1 2		

Appendix G: Pilot - item selection

For the pilot version of the TTT-NL, multiple adjustments and additions were made. A general question that was added to the score form is *Aan welk oor heeft u de telefoon?*[14] 'To which ear are you holding the phone?' as this might influence the perception of the verbal input (i.e. right ear advantage[15]).

Comprehension screening

Items opt2 and opt3, see Table G1, had been added as an alternative for item 2 as it only assumes two possible genders, which is not timeless and could cause problems in terms of inclusion in modern society. The alternative needed to have *nee* 'no' as an answer, as did the previous item.

Table G1. Proposed changes for the comprehension screening.

Comprehension screening		
Item	Question	Alternative question
2	*Bent u een man/vrouw?* 'Are you a male/female?'[16]	*opt2. Bent u een kind?* 'Are you a child?' *opt3. Bent u 18 jaar of ouder?* 'Are you 18 years old or above?'

Word finding

The original word finding task consisted of twenty items of which seventeen were nouns and three target words were verbs. The proportion of nouns and verbs is relatively off balance, which is why a new, equal division of nouns and verbs has been proposed. Furthermore, all other tests in the TTT-NL consisted of ten items at its maximum whereas the word finding test consisted of twenty items. Therefore, the aim was to construct a final, improved version of the word finding test that consists of ten test items and ten backup items in case there is doubt about the participant's ability to successfully complete the task. The items in Table G2 were changed because some of these items were outdated (i.e. *oma* 'grandma' and *krijt* 'chalk'), had less than an 80% rate of correctness (i.e. *wapens* 'weapons' and *puppy* 'puppy'), or the description could use improvement (i.e. *piloot* 'pilot'). Alternatives for these items have been found using SUBTLEX-NL (Keuleers et al., 2010), a database of Dutch word frequencies based on Dutch film and television subtitles. There had to be enough items in case some of the items would possibly excluded from the final version of the TTT-NL. The items had to be matched in frequency (i.e. same frequency as the eliminated item and similar frequencies for both sets of ten items), structure (i.e. open, gap-filling question, such as item 1 or closed question, such as item 6), and numbers (i.e. at least a ten-ten division of five nouns and five verbs). Some of these items were excluded from the pilot version of the TTT-NL. For the pilot items, eventually 31 items (seventeen nouns and fourteen verbs) have been included in the word finding test.

[14] Administered in the last four participants.

[15] It has been found that most listeners hear stimuli best in their right ear, which is most likely due to language being left-lateralized (Prete et al., 2018).

[16] The experimenter asked for the opposite gender, which means *nee* 'no' was de correct answer.

Table G2. Proposed changes for outdated items or items with less than an 80% rate of correctness.

Word finding[17]				
Item	Question	Target word	Alternative question	Target word
1.	*Kinderen worden vaak opgevangen door hun opa en …* 'Children are often looked after by their grandfather and …?'	*oma* 'grandmother'	*De man heeft een dochter, hij is dus een …* 'The man has a daughter, so he is a …'	*vader* 'father'
6.	*Hoe noem je pistolen, geweren en messen ook wel?* 'What do you call pistols, guns, and knives?'	*wapens* 'weapons'	*Hoe noem je meerdere soldaten bij elkaar?* 'What do you call multiple soldiers together?'	*leger* 'army'
13.	*Wie bestuurt er een vliegtuig?* 'Who flies an airplane?'	*piloot* 'pilot'	*Hoe noem je iemand die een vliegtuig bestuurt?* 'What do you call someone who flies an airplane?'	*piloot* 'pilot'
15.	*Waar schrijf je mee op een schoolbord?* 'What do you use to write on a chalkboard?'	*krijt* 'chalk'	*Hoe noem je een scheur in servies?* 'What do you call a crack in service?'	*barst* 'crack'
17.	*Hoe noem je een jong van een hond?* 'What do you call the young of a dog?'	*puppy* 'puppy'	*Waarmee is de vacht van een zebra bedekt?* 'With what is the fur of a zebra covered?'	*strepen* 'stripes'
19.	*Hoe heet een huis op het water?* 'What do you call a house on water?'	*woonboot* 'houseboat'	*Hoe noem je de plek waar schaatsers trainen?* 'What do you call the place where ice skaters train?' *Wat doet iemand die mensen in trance brengt?* 'What does someone who entrances people do?'	*ijsbaan* 'ice rink' *hypnotiseren* 'to hypnotise'
20.	*Hoe noem je mensen die dichtbij de noordpool wonen?* 'What do you call people who live close to the north pole?'	*eskimo/inuït* 'Eskimo/Inuit'	*Hoe noem je een kleine kubus met aan alle kanten een aantal ogen van 1 tot 6?* 'What do you call a small cube with on all sides a number of dots ranging from 1 to 6?' *Hoe noem je het wanneer een tuinman takken van een boom knipt?* 'What do you call it when a gardener cuts branches off a tree?'	*dobbelsteen* 'dice' *snoeien* 'to prune'

[17] Items 15, 17 and 20 ('dobbelsteen' *dice*) were excluded from the pilot version of the TeleTaalTest-NL.

Sentence completion

All items of the sentence completion task remained unchanged, except for item 10, see Table G3. The verb *hameren* 'to hammer away at' is not frequently used and is therefore considered to be too challenging and could elicit doubtful answers and/or hesitations.

Table G3. Proposed changes for the sentence completion task.

Sentence completion		
Item	Sentence	Alternative sentence
10.	*Het meisje hamert erop dat …* 'The girl hammers away at …'	*Het meisje benadrukt dat …* 'The girl emphasizes that …'

Appendix H: Results pilot - item selection

Comprehension screening

Item two was eventually replaced by opt2 *Bent u een kind?* 'Are you a child?' as this item matched the length of the original item most. All other items remained unchanged.

Word finding

After administering the pilot version, there remained a more or less equal set of pilot items with nouns or verbs as their target word. As noted earlier, the number of noun and verb target words was off balance and therefore, a new ten-ten division was made. The first ten items are the initial word finding test that consists of five verbs and five nouns. In case the test leader is uncertain about the patient's capacities with regard to word retrieval, the second set of items (again, five nouns and five verbs) can be administered. The number of items fits the length of all other subtests, excluding the comprehension screening, which makes the new TTT-NL a more balanced test with an even shorter duration than previous versions. While making the division, correctness (i.e. at least an 80% rate of correctness or above), word frequency based on the SUBTLEX-NL database (Keuleers et al., 2010), and word class (i.e. noun or verb) were taken into account. Additionally, items could not be preceded or followed by an item that was phonologically (i.e. *tang* - *wang* 'pincers - cheek') or semantically (i.e. *kat* - *hond* 'cat - dog') related to it in order to prevent the elicitation of phonemic or verbal (semantically related) paraphasias. Such items would make it hard to find out if the paraphasias, associated with certain subtypes of aphasia (Bastiaanse, 2011; Hallowell, 2017), are due to the test items or a possible aphasia.

Phonological repetition

The phonological repetition task was left intact, except for the second nonword *frimotika*. The consonant cluster at the beginning of the word was too difficult to get across by phone as the sound quality of fricatives (i.e. *f*, *v*), voiceless speech sounds (i.e. *f*, *s*), and speech sounds that have a closely related place of articulation (i.e. the labiodental *f* and bilabial *p* and *m*) are (partially) lost depending on the quality of the connection and the lack of visuals. Hence, *frimotika* was replaced by *prekamlédi*, an item from the DuLIP (De Witte et al., 2015) with a more or less similar word structure (i.e. consonant cluster and four syllables) but without the fricative at the beginning of the word.

Sentence completion

The sentence completion task was left unchanged. However, the description of *twijfel* 'doubt' in a participant's speech was too subjective and vague. Hence, the description has been extended and doubt is now expressed as *een pauze langer dan 4 seconden* 'a pause of more than 4 seconds'.