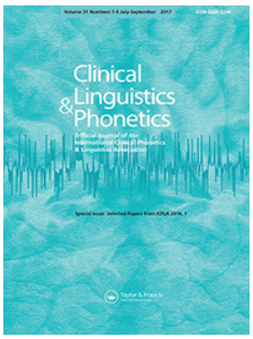

Clinical Linguistics & Phonetics

ISSN: 0269-9206 (Print) 1464-5076 (Online) Journal homepage: www.tandfonline.com/journals/iclp20

Cross-linguistic adaptations of *The Comprehensive Aphasia Test*: Challenges and solutions

Valantis Fyndanis, Marianne Lind, Spyridoula Varlokosta, Maria Kambanaros, Efstathia Soroli, Klaudia Ceder, Kleanthes K. Grohmann, Adrià Rofes, Hanne Gram Simonsen, Jovana Bjekić, Anna Gavarró, Jelena Kuvač Kraljević, Silvia Martínez-Ferreiro, Amaia Munarriz, Marie Pourquie, Jasmina Vuksanović, Lilla Zakariás & David Howard

To cite this article: Valantis Fyndanis, Marianne Lind, Spyridoula Varlokosta, Maria Kambanaros, Efstathia Soroli, Klaudia Ceder, Kleanthes K. Grohmann, Adrià Rofes, Hanne Gram Simonsen, Jovana Bjekić, Anna Gavarró, Jelena Kuvač Kraljević, Silvia Martínez-Ferreiro, Amaia Munarriz, Marie Pourquie, Jasmina Vuksanović, Lilla Zakariás & David Howard (2017) Cross-linguistic adaptations of *The Comprehensive Aphasia Test*: Challenges and solutions, Clinical Linguistics & Phonetics, 31:7-9, 697-710, DOI: 10.1080/02699206.2017.1310299

To link to this article: https://doi.org/10.1080/02699206.2017.1310299

Published with license by Taylor & Francis.© Valantis Fyndanis, Marianne Lind, Spyridoula Varlokosta, Maria Kambanaros, Efstathia Soroli, Klaudia Ceder, Kleanthes K. Grohmann, Adrià, Rofes, Hanne Gram Simonsen, Jovana Bjekić, Anna Gavarró, Jelena Kuvač Kraljević, Silvia Martínez-Ferreiro, Amaia Munarriz, Marie Pourquie, Jasmina Vuksanovií, Lilla Zakarićs, David Howard.

Published online: 27 Apr 2017.

Submit your article to this journal

Article views: 9803

View related articles

View Crossmark data

Citing articles: 27 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=iclp20

CLINICAL LINGUISTICS & PHONETICS
2017, VOL. 31, NOS. 7–9, 697–710
http://dx.doi.org/10.1080/02699206.2017.1310299

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS

Cross-linguistic adaptations of *The Comprehensive Aphasia Test*: Challenges and solutions

Valantis Fyndanis[a], Marianne Lind[a,b], Spyridoula Varlokosta[c], Maria Kambanaros[d], Efstathia Soroli[e], Klaudia Ceder[f], Kleanthes K. Grohmann[g], Adrià Rofes[h], Hanne Gram Simonsen[a], Jovana Bjekić[i], Anna Gavarró[j], Jelena Kuvač Kraljević[k], Silvia Martínez-Ferreiro[l], Amaia Munarriz[m], Marie Pourquie[n], Jasmina Vuksanović[o], Lilla Zakariás[p], and David Howard[q]

[a]Department of Linguistics and Scandinavian Studies/MultiLing (CoE), University of Oslo, Oslo, Norway; [b]Department of Speech and Language Disorders, Statped, Oslo, Norway; [c]Department of Linguistics, School of Philology, National and Kapodistrian University of Athens, Athens, Greece; [d]Department of Rehabilitation Sciences, Cyprus University of Technology, Limassol, Cyprus; [e]Département Sciences du Langage, University of Lille 3, Lille, France; [f]Institutionen för Neurovetenskap, Uppsala University, Uppsala, Sweden; [g]Department of English Studies, University of Cyprus, Nicosia, Cyprus; [h]Global Brain Health Institute, Trinity College Dublin, Dublin, Ireland; [i]Institute for Medical Research, University of Belgrade, Belgrade, Serbia; [j]Departament de Filologia Catalana, Universitat Autònoma de Barcelona, Bellaterra, Spain; [k]Department of Speech and Language Pathology, University of Zagreb, Zagreb, Croatia; [l]Department of Nordic Studies and Linguistics, University of Copenhagen, Copenhagen, Denmark; [m]Department of Linguistics and Basque Studies, University of the Basque Country, Vitoria-Gasteiz, Spain; [n]Basque Center on Cognition, Brain and Language, Donostia-San Sebastián, Spain; [o]Department of Philosophy Sciences, State University of Novi Pazar, Novi Pazar, Serbia; [p]Department of Linguistics, University of Potsdam, Potsdam, Germany; [q]School of Education, Communication and Language Sciences, Newcastle University, Newcastle upon Tyne, United Kingdom

ABSTRACT

Comparative research on aphasia and aphasia rehabilitation is challenged by the lack of comparable assessment tools across different languages. In English, a large array of tools is available, while in most other languages, the selection is more limited. Importantly, assessment tools are often simple translations and do not take into consideration specific linguistic and psycholinguistic parameters of the target languages. As a first step in meeting the needs for comparable assessment tools, the Comprehensive Aphasia Test is currently being adapted into a number of languages spoken in Europe. In this article, some key challenges encountered in the adaptation process and the solutions to ensure that the resulting assessment tools are linguistically and culturally equivalent, are proposed. Specifically, we focus on challenges and solutions related to the use of imageability, frequency, word length, spelling-to-sound regularity and sentence length and complexity as underlying properties in the selection of the testing material.

ARTICLE HISTORY
Received 3 December 2016
Revised 17 March 2017
Accepted 20 March 2017

KEYWORDS
Aphasia; assessment; Comprehensive Aphasia Test (CAT); cross-linguistic adaptations; outcome measures

Introduction

Aphasia is characterised by extensive variation in severity and type of language impairment as well as in management strategies and impact on quality of life, both within and across speakers with aphasia and their families (Whitworth, Webster, & Howard, 2014). Moreover, aphasia

CONTACT Spyridoula Varlokosta svarlokosta@phil.uoa.gr National and Kapodistrian University of Athens, Department of Linguistics, Faculty of Philology, School of Philosophy, Panepistimioupoli 15784, Zografou, Athens, Greece.

© Valantis Fyndanis, Marianne Lind, Spyridoula Varlokosta, Maria Kambanaros, Efstathia Soroli, Klaudia Ceder, Kleanthes K. Grohmann, Adrià, Rofes, Hanne Gram Simonsen, Jovana Bjekić, Anna Gavarró, Jelena Kuvač Kraljević, Silvia Martínez-Ferreiro, Amaia Munarriz, Marie Pourquie, Jasmina Vuksanovií, Lilla Zakarićs, David Howard. Published with license by Taylor & Francis.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

research is heterogeneous and characterised by language-, country- and discipline-specific features (e.g. Gitterman, Goral, & Obler, 2012; Menn & Obler, 1990). Hence, comparison across borders (linguistic, regional or disciplinary) is often difficult or impossible to achieve. A particular obstacle to comparative research on aphasia and aphasia rehabilitation is the lack of comparable assessment tools and outcome measures across different languages. This diminishes the potentially cumulative strength and broader relevance of the evidence (cf. Collaboration of Aphasia Trialists, 2013, and references therein).

With respect to English, there are many — primarily diagnostic — tools available, such as the *Boston Diagnostic Aphasia Examination* (Goodglass & Kaplan, 1972, 1983; Goodglass, Kaplan, & Barresi, 2001), the *Western Aphasia Battery* (Kertesz, 1982), the *Comprehensive Aphasia Test* (Swinburn, Porter, & Howard, 2004), the *Bilingual Aphasia Test* (Paradis & Libben, 1987), the *Aachen Aphasia Test* (Miller, De Bleser, & Willmes, 1997; Miller, Willmes, & De Bleser, 2000), the *Porch Index of Communicative Ability* (Porch, 1971) or the *Psycholinguistic Assessment of Language Processing in Aphasia* (Kay, Lesser, & Coltheart, 1992), to name a few. There are also several tools focusing on particular (linguistic, psychosocial, etc.) aspects of aphasia. In most other languages, the selection is scarce, and often the tools are direct translations of one or more of the above-mentioned tests (see also Ivanova & Hallowell, 2013). In fact, the lack of adapted and clinically feasible comprehensive cross-linguistic assessment tools hinders not only comparability of aphasia assessment results across languages for both clinical and research purposes, but also the management of multilingual individuals with post-stroke aphasia.

The *Collaboration of Aphasia Trialists* (COST Action IS1208, 2013–2017) — which is an EU-funded network of multidisciplinary aphasia investigators from 26 countries, mainly in Europe — acknowledges the need for comparable assessment tools across languages spoken in Europe, adapted to the specificities and the linguistic properties of the languages. As a first step in meeting this need, *The Comprehensive Aphasia Test (CAT)*, originally developed for English by Swinburn, Porter, and Howard (2004), is currently being adapted into the following languages which are spoken in Europe: Basque, Catalan, Croatian, Cypriot Greek, French, (Standard Modern) Greek, Hungarian, Norwegian, Serbian, Spanish, Swedish and Turkish.[1] The present methodological article is based on the CAT adaptations into these 12 languages. Such an adaptation enterprise allows us to pave the way for collecting and pooling large amounts of comparable data with different versions of the same outcome measure and facilitate more robust research on aphasia rehabilitation (see also Ali, English, Bernhardt, Sunnerhagen, & Brady, 2013; Brady, Ali, Fyndanis, Kambanaros, Grohmann, Laska, Hernández-Sacristán, & Varlokosta, 2014). The present article is based on the collaborative effort carried out by researchers and clinicians who are members of the second working group (WG2) of the *Collaboration of Aphasia Trialists*,[2] which specifically addresses questions on assessment and outcomes.

Rationale for adapting the CAT

Once acknowledging the need for comparable cross-linguistic assessment tools, the first step of the process was to define the criteria that such a test should meet. Previous research (Tsapkini, Vlachou, & Potagas, 2010; Ivanova & Hallowell, 2013 among others) has stressed the fact that a cross-linguistic tool should be comprehensive, relatively short and clinically relevant (in helping to design treatment approaches and inform further assessment). Furthermore, it should be user friendly in order to accommodate different professionals involved in the assessment process (such as speech and language therapists,

medical doctors, neuropsychologists or neurolinguists, as is often the case in many countries across Europe). The final tool should then ideally be standardised, normed and validated for each of the target languages.

The second step was to review existing aphasia assessment tools, evaluate to what extent there is a need to develop new ones or eventually adapt one of the existing tools to the languages that lack such a test. Several commonly used assessment tools were evaluated. The most striking realisation was that there is a lot of variability in the domain of assessment across countries. Of major concern was the fact that a number of countries have no assessment tools available, as there is no diagnostic tradition for aphasia in place (e.g. Croatia or Cyprus), nor do they have tools suited to their sociolinguistic reality (e.g. in the Basque Autonomous Community, aphasia assessment tools are only available in Spanish, despite the fact that both Spanish and Basque are official languages). We concluded that there was no need to create a new tool and that the CAT (Swinburn et al., 2004) met the criteria for the intended cross-linguistic use.[3]

The CAT is comprehensive in its scope, as it covers not only the purely linguistic aspects of aphasia, but it also includes a cognitive screening as well as a questionnaire focusing on disability and quality of life. Despite its comprehensiveness, the CAT is a relatively short test, which is also normed and validated for English, and widely used in the English-speaking world. Thus, it can be used both as an assessment tool that will provide the clinician with a summary of the aphasic speaker's linguistic abilities and impairments, on the basis of which language remediation can be designed, and further 'as a means of monitoring recovery and measuring outcome' (Howard, Swinburn, & Porter, 2010a: 59). This methodological article focuses on the language battery of the CAT, since it is the most challenging component due to its inherent cross-linguistic and cross-cultural variability.

Following Paradis and Libben (1987) and Bates, Wulfeck, and MacWhinney (1991) on the inappropriateness of direct or literal translations of existing assessment tools between languages, we decided to develop adaptations of the CAT and not simple translations.[4] Previous research underlines the need to take into account not only the linguistic but also the cultural differences between language communities (Edwards & Bastiaanse, 2007; Roberts, 2008, among others) when choosing to adapt an existing assessment tool to a particular language. Thus, the developed adaptations presented here take into account both linguistic and cultural differences across systems and language communities.

With respect to the language battery, the tool involves one comprehension part (subtests 7 to 11) and one expressive language part (subtests 12 to 27) that cover a wide range of areas and that are 'designed to elicit as much information as possible about the factors determining performance accuracy in each task within a limited number of items' (Howard et al., 2010a: 60):

(1) auditory comprehension: single word (subtest 7), sentence (subtest 9), and paragraph level (subtest 11);
(2) written comprehension: single word (subtest 8) and sentence level (subtest 10);
(3) repetition: simple words (subtest 12), complex words (subtest 13), non-words (subtest 14), digit strings (subtest 15), and sentences (subtest 16);
(4) verbal expression: objects (subtest 17), actions (subtest 18), and picture description (subtest 19);

(5) oral reading: including regular and irregular, short and long simple words (subtest 20), complex words (subtest 21), function words (subtest 22), and non-words (subtest 23);
(6) written expression: same- and cross-case letter and word copying (subtest 24), single word confrontation naming (subtest 25), writing to dictation (subtest 26), and picture description (subtest 27).

Importantly, unlike many other assessment tools, the linguistic properties and psycholinguistic variables underlying the choice of subtests and materials in the CAT are stated *explicitly* in the manual, which greatly facilitates the adaptation procedure. These include word frequency, imageability, word length, phonological and semantic relatedness, orthographic regularity, animacy, morphological complexity, sentence length and syntactic complexity.[5] These variables have all been shown to affect performance of people with aphasia on a variety of tasks (e.g. Bastiaanse, Bouma, & Post, 2009; Bastiaanse, Wieling, & Wolthuis, 2015; Whitworth et al., 2014, and references therein). Below we underline the need to address challenges related to the adaptation of the testing material with respect to these variables and the solutions proposed.

The CAT adaptation process

A key feature of the adaptation process was to select adequate material on the basis of the linguistic properties (word length, spelling-to-sound regularity, sentence length and complexity) and psycholinguistic variables (imageability and frequency) underlying the test items of the original CAT, but also to address certain problems related to cultural differences across language communities and to some limitations of the test.

For each language, local teams were formed. Each language team followed a committee approach (Hambleton, 2005), actively involving stakeholders including speech and language therapists, linguists, and in some cases psychologists and individuals with aphasia in the process. The adaptations of the CAT into the specific languages followed the guidelines that had been jointly established and agreed upon during the regular meetings of WG2 'Assessment and Outcomes' within COST Action IS1208.

One of the first decisions was not to change the basic format of the test in order to maximise comparability of versions across languages. This means that each language version of the CAT had to have the same number of subtests and items in each subtest, and adhere to the original scoring system.

We also decided to adapt rather than translate the test. Thus, we kept the same underlying linguistic and psycholinguistic variables in all the new versions as in the original CAT, while respecting cultural differences across the different language communities. Many of the languages involved in the project lacked the necessary background data on variables such as frequency and imageability. Hence, as part of the adaptation process, these data had to be collected. Moreover, there were many new items in the language battery that had to be created in order to replace those of the original version that were not adequate or did not respect the linguistic properties and/or cultural constraints of the particular language communities. Consequently, a large number of new illustrations had to be drawn. The new illustrations for the adaptations needed to be as similar as possible to the original (i.e. black and white line drawings), stylistically close and culturally appropriate.

Challenges and solutions

All the adaptations were adaptations of the English CAT into different languages. Challenges included: (a) cultural differences across language communities; (b) between-language community differences regarding the availability of background data on variables such as frequency and imageability; and (c) specific cross-linguistic differences involving different language levels, such as phonology (e.g. spelling-to-sound regularity), syntax (e.g. sentence complexity) and the lexicon (e.g. word length).

a. Cultural appropriateness

With respect to the cultural aspect, some items of the original test were inappropriate for different reasons. This was taken into serious consideration in the different adaptations. For example, the main characters and actions of some items in the English CAT (e.g. 'butcher', 'nun', 'winding a watch', 'licking a stamp') were not used in some languages, as these are obsolete or not equally familiar across all the language communities. Another example was the typical representation of items like 'school' and 'church', which may be different across cultures. Therefore, appropriate pictures were selected for each language community, based on a consensus among the members of the adaptation teams and/or on a naming agreement test. It is important to note that verbs with a negative connotation (e.g. 'killing', 'shooting') were avoided in all languages.

While respecting the structural and cultural differences across languages, we tried to make the design of each of the tests as similar to the original as possible. For example, for the sentence comprehension subtests, the same picture choice was used for all languages, as developed by consensus, although the sentences used with those pictures varied between languages depending on the structural properties of the particular language system (for more details on the linguistic adaptation, see section c).

b. Underlying psycholinguistic variables

The selected psycholinguistic variables that were taken into account in the present CAT adaptation are imageability and word frequency.

Imageability

Imageability is a psycholinguistic variable that is used to explain the relative ease with which a word gives rise to a mental image or sensory experience (Paivio, Yuille, & Madigan, 1968). Imageability values are established on the basis of native speakers' ratings. For some of the languages, and prior to this work, imageability values had been established for a wide range of words (e.g. Bird, Franklin, & Howard (2001) and Cortese & Fugett (2004) for English; Simonsen, Lind, Hansen, Holm, & Mevik (2013) for Norwegian; Desrochers & Thompson (2009) for French). However, this was not the case for many of the other languages into which the CAT is being adapted (e.g. Basque, the two varieties of Greek, Hungarian or Serbian).

For languages without established imageability values, imageability ratings were collected for the relevant candidate words to be included in the CAT.[6] This was done through the use of questionnaires, for which at least 20 participants rated each selected word on a

7-point Likert scale, with one representing 'impossible to create an image' and seven representing 'very imageable'. When selecting items for these subtests, it was of course important to ensure that there was a gap between the high and low imageability items, respectively.

The contrast between high and low imageability was relevant for word repetition (subtest 12), word reading (subtest 20), object naming (subtest 17), written picture naming (subtest 25) and writing to dictation (subtest 26).

Frequency

The second property included was frequency. Word frequency is a psycholinguistic variable that is used to indicate how often a word is used, be it in spoken or written form (Bastiaanse et al., 2015). Ideally, frequency measures should be taken from spoken corpora (Brysbaert & New, 2009), as was done in the CAT adaptation for French. Written corpora may also be used, as written and spoken frequency corpora typically correlate highly (e.g. Pastizzo & Carbone, 2007). Hence, web-based or other written corpora as a proxy for spoken language were used as an alternative to collect ratings for this variable (Basque, Catalan, Cypriot Greek, Hungarian, Norwegian, Greek, Swedish). In languages without such corpora, familiarity ratings can be collected (i.e. subjective ratings of how often one uses the word; Noble, 1953), as was done for Croatian and Turkish. It has been found that frequency and familiarity scores significantly correlate (see Tanaka-Ishii & Terada, 2011, and references therein). Moreover, a combination of corpora and familiarity ratings was used for Serbian and Spanish.

The language battery of the CAT incorporates a contrast between high- and low-frequency words in tests involving comprehension of spoken words (subtest 7), word repetition (subtest 12), object naming (subtest 17), written picture naming (subtest 25) and writing to dictation (subtest 26). In order to find appropriate words for the relevant adaptations, one needs to decide on which frequencies are to be considered as low and which ones as high. However, the distribution of word-frequency values is heavily right-skewed with a lot of low-frequency words and relatively few high-frequency words, making it difficult to statistically determine the low and high frequency cut-offs. Not every language adopted the same cut-off criteria, but as a common principle, there was a gap between high- and low-frequency exemplars.[7]

c. Linguistic properties

Some of the linguistic challenges stemmed from the fact that languages largely differ in their typological properties (with respect to phonology, semantics, (morpho)syntax, lexicon, etc.), and this is the case for both closely related (e.g. Norwegian vs. English) and more distantly related languages (e.g. Turkish vs. English or Basque vs. English). For example, English orthography is less transparent than the orthography of languages such as Greek and Spanish (e.g. Seymour, Aro, & Erskine, 2003). With respect to sentence structure, passive sentences are more commonly used in English than in languages such as Basque, Croatian and Greek. These are only some cases where matching between different languages was challenging. Table 1 presents the selected linguistic properties that were taken into account in the present CAT adaptation, as accommodated by each language team.

Table 1. Selected linguistic properties taken into account in the CAT and the way they were accommodated by each language team.

	Orthographic regularity	Word length	Sentence complexity (complex sentences used)
Basque	Implemented	Disyllabic & tetrasyllabic words	Simple active derived (OVS) sentences
Catalan	Dropped as irrelevant	Monosyllabic, disyllabic, trisyllabic, & tetrasyllabic	Passives
Croatian	Implemented	Disyllabic & tetrasyllabic words	Active sentences with object-extracted relative clauses and active simple derived (OVS) sentences
Cypriot Greek	Implemented with a compromise	Disyllabic & tetrasyllabic words	Active sentences with object-extracted relative clauses
French	Implemented	Monosyllabic & trisyllabic words	Passives and object-extracted relative clauses
Greek	Implemented with a compromise	Disyllabic & tetrasyllabic words	Active sentences with object-extracted relative clauses
Hungarian	Implemented with a compromise	Word comprehension and naming: Monosyllabic & trisyllabic Repetition of words: Disyllabic & tetrasyllabic words	Active sentences with object-extracted relative clauses and active simple derived (OVS) sentences
Norwegian	Implemented	Monosyllabic & trisyllabic words	Passives
Serbian	Dropped as irrelevant		Active sentences with object-extracted relative clauses and active simple derived (OVS) sentences
Spanish	Implemented	Disyllabic & tetrasyllabic words	Passives
Swedish	Implemented	Monosyllabic & trisyllabic words	Passives
Turkish	Implemented	Monosyllabic & trisyllabic words	Active sentences with object-extracted relative clauses

Spelling-to-sound regularity

Spelling-to-sound regularity refers to the relative ease with which the phonetic realisation of a word can be predicted on the basis of its orthography/spelling and, conversely, to the relative ease with which the orthography of a word can be predicted on the basis of its phonetic realisation. A word is considered as 'regular' when its pronunciation or spelling is correctly produced by the grapheme-sound correspondence rules of a language and 'irregular' when its pronunciation or spelling cannot be predicted from these rules (Protopapas & Vlahou, 2009; Ziegler, Perry, & Coltheart, 2003). The relevant subtests that involved this variable are reading words (subtest 20), written picture naming (subtest 25) and writing to dictation (subtest 26).

Languages vary in the degree of this type of regularity. In languages with opaque orthography, such as English and French, there are many irregular words (see Ziegler, Perry, & Coltheart, 2000). For example, the word 'yacht' in English is pronounced [ˈjɒt], not [jætʃt], and, conversely, the spelling/orthography of this word cannot be predicted on the basis of its pronunciation ([ˈjɒt]). Additionally, while some items involved irregular words in the original CAT (e.g. *giraffe*), their equivalent translation would provide a regular item (e.g. in French: *girafe*) and thus would not be adequate for the adaptation. For languages with transparent orthographies (e.g. Basque, Greek, Serbian, Spanish, Turkish), a contrast between regular and irregular words is not available. One solution to test for regularity errors in such languages was to include words for spelling or reading where at least one sound unit corresponds to more than one grapheme (e.g. Greek: the word *αίμα* /*éma*/'blood', where the combination of the first two graphemes <a> and <i> is pronounced as the single [e]). Given that in languages like Greek some sounds have more

than one possible orthographic realisation (for instance, [e] can be realised as <aɪ> or as <ɛ>), the orthography of words including such sounds is unpredictable on the basis of their phonetic realisation. Other languages with transparent orthography (e.g. Basque) opted instead simply to drop this as a variable in reading.

Word length

Another linguistic property of the CAT is word length. In the original CAT, mono- and tri-syllabic words (e.g. head, hospital) were included as short and long words, respectively, in word repetition (subtest 12), object naming (subtest 17), reading words (subtest 20), written picture naming (subtest 25) and writing to dictation (subtest 26). However, what is long and what is short differs across languages. English, for example, has many monosyllabic words, but morphologically rich languages such as Basque, Greek, Hungarian, Spanish and Turkish have few monosyllabic words.

Given this difference between languages, we decided that languages with few monosyllabic words should include di- and tetra-syllabic words as short and long words, respectively (e.g. Spanish) (Table 1). What was important was to keep the one-syllable gap between short and long words (i.e. successively from one- to four-syllable words). Short and long words differed both in number of syllables and sounds.

Sentence length and complexity

A final major challenge relating to linguistic structure is the property of sentence length and complexity relevant for comprehension of spoken sentences (subtest 9), comprehension of written sentences (subtest 10), comprehension of paragraphs (subtest 11) and repetition of sentences (subtest 16).

For sentence length, both content and function words were relevant. However, given that languages differ in a number of morphosyntactic variables (free-standing function words, use of pronouns, temporal and aspectual marking, etc.), adapted sentences could not always consist of the same number of words across languages. Hence, in the sentence repetition subtest, we decided to keep the number of content words constant but not necessarily the same overall number of words. This yields, for example, the English sentence 'They decided to paint the room blue' (four content words, seven words in total) and its Swedish equivalent 'De valde att måla rummet blått' (four content words, six words in total) or its French equivalent 'Ils ont decidé de peindre la chambre en bleu' (four content words, but nine words in total). Other adaptations included the replacement of the progressive present used in the English version by simple present forms (for instance, in Basque, French, Greek, Norwegian, Serbian and Spanish), the addition of the omitted pronoun when relevant (e.g. English: The carpet the cat is on is green; French: Le tapis, où le chat se repose est vert), and adjective-noun inversions (e.g. English: The green cat is on the carpet; French: Le chat vert est sur le tapis 'The cat green is on the carpet').

Regarding structural complexity, passive sentences are included in the original CAT as a type of structurally complex sentences. However, while passive sentences are frequently used in English, they are either relatively infrequent in languages such as Croatian, Greek and Serbian or non-existent, like in Basque. In these languages, passives cannot be used as an appropriate example of structurally complex sentences. We therefore decided that for such languages, other structurally complex sentences should be used, for example,

sentences with object-extracted relative clauses (e.g. Greek) or with derived word order (OVS) (e.g. Basque) (Table 1). These sentence types are similar in that they are non-canonical and involve syntactic displacement. For examples of passive sentences, sentences with object-extracted relative clauses and OVS sentences, see (1), (2) and (3), respectively.

(1) The boy$_i$ was kissed __$_i$ by the girl.
(2) The boy$_i$ that the girl kissed __$_i$ is tall.
(3) Mutila-ø$_i$ besarkatu zuen$_k$ neska-k __$_k$ __$_i$. (Basque)

boy-ABS kiss AUX girl-ERG
'The boy was kissed by the girl'

Discussion

This methodological article was motivated by the need for common assessment and outcome measures across different languages, as identified within COST Action IS1208, *Collaboration of Aphasia Trialists*. This is crucial for aphasia research and clinical practice. In addition, it allows for comparisons between groups of monolingual patients from different language communities. It will also facilitate the assessment of different languages spoken by multilingual individuals with aphasia.

The Comprehensive Aphasia Test (CAT; Swinburn et al., 2004) was adapted into a number of languages. This test is comprehensive in that it assesses both the language impairment in the context of cognitive difficulties and the impact of communication impairment on the participant's quality of life. Most importantly, the CAT is also comprehensive in the sense that its language battery covers all language modalities and levels of linguistic analysis. The adaptation of the CAT represents an update of the state of the art in aphasia research and practice in many countries, as it facilitates the adoption of the cognitive neuropsychological approach to assessment and rehabilitation of language (Whitworth et al., 2014). In some cases, the CAT adaptation is the first development of a properly designed aphasia test for a language (e.g. Croatian, Cypriot Greek).

Adopting the same design principles as in the original CAT, versions of the test were developed in a variety of typologically different languages keeping the same test format and respecting the linguistic and cultural differences across different language communities. This is very different from past developments of tests in other languages, which often depended on literal translation of the materials.

A limitation we acknowledge stems from the fact that some properties have different degrees of importance across languages. For example, morphological complexity is more relevant in highly inflected languages such as Basque, Greek and Spanish, and spelling-sound regularity is more important in orthographically non-transparent languages such as English and French. Since English is not a highly inflected language, morphological complexity was hardly assessed in the original CAT. This was maintained in the adaptations in order to keep comparability. For all the other challenges stemming from the differences across languages and cultures, acceptable solutions were found. The major challenge was to find items and structures related to the underlying properties that are comparably equivalent across languages.

Concluding, the advantages of adapting an existing aphasia assessment tool into a number of different languages outweigh the limitations and challenges encountered. This is so because basing the adaptations on the linguistic and psycholinguistic properties that are central to the original tool helps in constructing equivalent adaptations across languages. Moreover, keeping the same number of subtests and items as well as the same scoring criteria facilitates comparison between languages. It also opens up for larger scale investigations of aphasia.

One of the goals of the present article is to offer this test as an example for adaptations in other languages and to seek to support clinicians and researchers who wish to develop versions of the test in other languages. Future plans include standardisation, norming and validation of the new adaptations of the CAT using data from people with aphasia and neurologically healthy individuals. The validity of the CAT will thus be established for a number of different purposes, including assessment of people with aphasia, measuring change over time, and as an outcome measure for aphasia trials.

Acknowledgments

The work reported in this article was undertaken within Working Group 2 (WG2) (Assessment and Outcomes) of COST Action IS1208, *Collaboration of Aphasia Trialists* (PI and Chair: Marian Brady). Spyridoula Varlokosta was the Lead of WG2 from August 2014 to April 2017 and served as a Deputy Lead from May 2013 to July 2014. Valantis Fyndanis was the Deputy Lead of WG2 from August 2014 to April 2017. We would like to thank Roelien Bastiaanse for leading WG2 from May 2013 to July 2014 and for motivating us to undertake the CAT adaptation project. All authors substantially contributed to the meetings of WG2, where the guidelines for the adaptation process of the CAT were proposed and discussed. The first nine authors and the last author contributed to the writing of the article and to the revisions performed in the review process. The remaining authors provided substantial comments on the original and revised articles. We also thank the following colleagues for their contributions to the work undertaken within WG2: Ingrida Balčiūnienė, Charlotte Jacquemot, Line Haaland-Johansen, Carlos Hernández Sacristán, Monica Blom Johansson, Melita Kovacevic, İlknur Maviş, Carolina Méndez-Orellana, Ingvild Røste, Io Salmons, Ingrid Sör and Müge Tunçer.

Declaration of interest

The authors report no conflicts of interest.

Funding

For the work reported in this article, we were supported by various funding bodies. Our meetings in Working Group 2 were partly supported by grants from COST (European Cooperation in Science and Technology) Action IS1208, *Collaboration of Aphasia Trialists*. The first author was supported by a Short-Term Scientific Mission grant (reference number COST-STSM-ECOST-STSM-IS1208-131116-080878). The Basque adaptation was partly supported by the MINECO/FEDER (FFI2015-68589-C2-1-P) and the Basque Government (IT983-16 – GIC 15/129). The Catalan adaptation was partly supported by project FFI2014-56968-C4-1-P. The Croatian adaptation was supported by the Croatian Science Foundation through the project *Adult language processing* (ALP, HRZZ-2421) and by the Operational Programme Human Resources Development through the project *Interdisciplinary approach to language model of dyslexia in adult* (HR.3.2.01-0247 EU-ESF). The French adaptation was partly supported by Agence Nationale de la Recherche (ANR) grants (#ANR-10- LABX-0087 and #ANR-

10-IDEX-0001-02). The Norwegian adaptation was partly supported by the Research Council of Norway through its Centres of Excellence funding scheme, project number 223265. The Serbian adaptation was supported by a project grant (#175012) from the Ministry for Education, Science and Technological Development of the Republic of Serbia. The Spanish adaptation was partly supported by PROGRAM (University of Copenhagen Excellence Programme for Interdisciplinary Research) and projects FFI2015-68589-C2-1-P and FFI2014-61888-EXP (Ministerio de Economía y Competitividad, Spain). The Swedish adaptation was funded by the Swedish National Aphasia Foundation. The Turkish adaptation was partly supported by the Anadolu University Scientific Research Projects Board, project number 1509S632.

Notes

1. Prior to the initiative of our Collaboration, the CAT was adapted to Danish (Swinburn, Porter, & Howard, 2014) and Dutch (Visch-Brink, Vandenborre, de Smet, & Mariën, 2014). We understand that versions of this test are currently in development for Gulf Arabic, Japanese, Mandarin Chinese, and other languages.
2. Our collaborative network is divided into five WGs. More information on this COST Action can be found here: http://www.aphasiatrials.org/
3. The *Bilingual Aphasia Test* (Paradis & Libben, 1987) has been developed for an impressive range of languages and could have been a candidate for our purposes. However, although it is comprehensive in the sense that it covers a wide range of linguistic domains, it is limited in that it only assesses language. Furthermore, it is long for an assessment tool in a clinical setting, and it is only normed for very few of the languages.
4. Note that several aphasia tests are currently used in different language communities that constitute translations, not adaptations (e.g., the *Boston Naming Test*; for French: Lapointe-Goupil, Everett, Rousseau, Giguère, Laplante, & Keller, 2004; Demers, Robillard, Lafleche, Nash, Heyman, & Fillenbaum, 1994; for Swedish: Tallberg, 2005; for Spanish: Allegri, Villavicencio, Taragano, Rymberg, Mangone, & Baumann, 1997). Given our methodological choice, techniques such as forward and backward translation that are often used in translations of questionnaires are not applicable to our work (cf. the World Health Organization's recommendations at http://www.who.int/substance_abuse/research_tools/translation/en).
5. For more details about the strengths and weaknesses of the CAT, see Howard, Swinburn, & Porter (2010a, 2010b), Bruce & Edmundson (2010), and Springer & Mantey (2010), among others.
6. Rofes, Zakariás, Ceder, Lind, Bloom Johansson, Bjekić, Fyndanis, Gavarró, Simonsen, Hernández-Sacristán, Kuvač Kraljević, Martínez-Ferreiro, Mavis, Méndez Orellana, Meteyard, Salmons, Sör, Tunçer, Vuksanovic, Varlokosta, & Howard (2016) found significant and moderate to strong correlations between the imageability ratings of English words and ratings collected for most of the languages under consideration.
7. See Van Heuven, Mandera, Keuleers, & Brysbaert (2014) for an algorithm implementing Zipf's law on the distribution of frequency values (Zipf, 1949).

References

Allegri, R. F., Villavicencio, A. F., Taragano, F. E., Rymberg, S., Mangone, C. A., & Baumann, D. (1997). Spanish boston naming test norms. *The Clinical Neuropsychologist*, *11*, 416–420.

Ali, M., English, C., Bernhardt, J., Sunnerhagen, K. S., & Brady, M. (2013). More outcomes than trials: A call for consistent data collection across stroke rehabilitation trials. *International Journal of Stroke*, *8*, 18–24.

Bastiaanse, R., Bouma, G., & Post, W. (2009). Linguistic complexity and frequency in agrammatic speech production. *Brain and Language*, *109*, 18–28.

Bastiaanse, R., Wieling, M., & Wolthuis, N. (2015). The role of frequency in the retrieval of nouns and verbs in aphasia. *Aphasiology*, *30*, 1–18.

Bates, E., Wulfeck, B., & MacWhinney, B. (1991). Cross-linguistic research in aphasia: An overview. *Brain Language*, *41*, 123–148.

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers*, *33*, 73–79.

Brady, M. C., Ali, M., Fyndanis, C., Kambanaros, M., Grohmann, K. K., Laska, A. C… Varlokosta, S., on behalf of the Collaboration of Aphasia Trialists (2014). Time for a step change? Improving the efficiency, relevance, reliability, validity and transparency of aphasia rehabilitation research through core outcome measures, a common data set and improved reporting criteria. *Aphasiology*, *28*, 1385–1392.

Bruce, C., & Edmundson, A. (2010). Letting the CAT out of the bag: A review of the comprehensive aphasia test. Commentary on Howard, Swinburn, and Porter, "putting the CAT out: What the comprehensive aphasia test has to offer". *Aphasiology*, *24*, 79–93.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*, 977–990.

Collaboration of Aphasia Trialists (2013). *Memorandum of understanding*. Retrieved May 2016 from http://www.cost.eu/COST_Actions/isch/IS1208 http://w3.cost.eu/fileadmin/domain_files/ISCH/Action_IS1208/mou/IS1208-e.pdf

Cortese, M. J., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words. *Behavior Research Methods, Instruments, & Computers*, *36*, 384–387.

Demers, P., Robillard, A., Lafleche, G., Nash, F., Heyman, A., & Fillenbaum, G. (1994). Translation of clinical and neuropsychological instruments into French: The CERAD experience. *Age and Ageing*, *23*, 449–451.

Desrochers, A., & Thompson, G. L. (2009) Subjective frequency and imageability ratings for 3600 French nouns. *Behavior Research Methods*, *41*, 546–557.

Edwards, S., & Bastiaanse, R. (2007). Assessment of aphasia in a multi-lingual world. In M. J. Ball & J. S. Damico (Eds.), *Clinical aphasiology: Future directions* (pp. 245–258). New York, NY: Psychology Press.

Gitterman, M., Goral, M., & Obler, L. K. (Eds.) (2012). *Aspects of multilingual aphasia*. Clevedon, UK: Multilingual Matters.

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

Goodglass, H., & Kaplan, E. (1983). *The assessment of aphasia and related disorders* (2nd *edition)*. Philadelphia, PA: Lea & Febiger.

Goodglass, H., Kaplan, E., & Barresi, B. (2001). *The assessment of aphasia and related disorders* (3rd *ed.)*. Boston, MA: Lippincott Williams & Wilkins.

Hambleton, R. K. (2005). Issues, designs, and technical guidelines for adapting tests into multiple languages and cultures. In R. K. Hambleton, P. F. Merenda, & C. D. Spielberger (Eds.), *Adapting educational and psychological tests for cross-cultural purposes* (pp. 3–38). Mahwah, NJ: Lawrence Erlbaum.

Howard, D., Swinburn, K., & Porter, G. (2010a). Putting the CAT out: What the comprehensive aphasia test has to offer. *Aphasiology*, *24*, 56–74.

Howard, D., Swinburn, K., & Porter, G. (2010b). Putting the CAT out: A response to the commentaries. *Aphasiology*, *24*, 94–98.

Ivanova, M., & Hallowell, B. (2013). A tutorial on aphasia test development in any language: Key substantive and psychometric considerations. *Aphasiology*, *27*, 891–920.

Kay, J., Lesser, R., & Coltheart, M. (1992). *Psycholinguistic Assessments of Language Processing in Aphasia*. Hove, UK: Lawrence Erlbaum.

Kertesz, A. (1982). *The Western Aphasia Battery*. New York, NY: Grune & Stratton.

Lapointe-Goupil, R., Everett, J., Rousseau, F., Giguère, V., Laplante, L., & Keller, E. (2004). Étude comparative entre le Boston Naming Test et le Test de Dénomination Bellefleur. *Revue québécoise de psychologie*, *25*, 203–212.

Menn, L., & Obler, L. K. (1990). Cross-language data and theories of agrammatism. In L. Menn & L. K. Obler (Eds.), *Agrammatic aphasia* (vol. 2, pp. 1369–1389). Amsterdam, The Netherlands: John Benjamins.

Miller, N., De Bleser, R., & Willmes, K. (1997). The English language version of the Aachen Aphasia Test. In W. Zeiegler & K. Deger (Eds.), *Clinical phonetics and linguistics* (pp. 257–265). London, UK: Whurr.

Miller, N., Willmes, K., & De Bleser, R. (2000). The psychometric properties of the English language version of the Aachen Aphasia Test (EAAT). *Aphasiology*, *14*, 683–722.

Noble, C. E. (1953). The meaning-familiarity relationship. *Psychological Review*, *60*, 89–98.

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, *76*, 1–25.

Paradis, M., & Libben, G. (1987). *The assessment of bilingual aphasia*. Hillsdale, NJ: Lawrence Erlbaum.

Pastizzo, M. J., & Carbone, R. F. (2007). Spoken word frequency counts based on 1.6 million words in American English. *Behavior Research Methods*, *39*, 1025–1028.

Porch, B. E. (1971). *Porch Index of Communicative Ability*. Palo Alto, CA: Consulting Psychologists Press.

Protopapas, A., & Vlahou, E. L. (2009). A comparative quantitative analysis of Greek orthographic transparency. *Behavior Research Methods*, *41*, 991–1008.

Roberts, P. M. (2008). Issues in assessment and treatment for bilingual and culturally diverse patients. In R. Chapey (Ed.), *Language intervention strategies in aphasia and related neurogenic communication disorders* (5th *ed.)* (pp. 245–275). Philadelphia, PA: Lippincott Williams & Wilkins.

Rofes, A., Zakariás, L., Ceder, K., Lind, M., Bloom Johansson, M., Bjekić, J… Howard, D. (2016). Word imageability from a cross-linguistic perspective. *Stem-, Spraak- en Taalpathologie, 21*, 158–161.

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*, 143–174.

Simonsen, H. G., Lind, M., Hansen, P., Holm, E., & Mevik, B.-H. (2013). Imageability of Norwegian nouns, verbs and adjectives in a cross-linguistic perspective. *Clinical Linguistics & Phonetics*, *27*, 435–446.

Springer, L., & Mantey, S. (2010). The Comprehensive aphasia test: A review. commentary on Howard, Swinburn, and Porter. "Putting the CAT out: What the comprehensive aphasia test has to offer". *Aphasiology*, *24*, 75–78.

Swinburn, K., Porter, G., & Howard, D. (2004). *Comprehensive Aphasia Test*. Hove, UK: Psychology Press.

Swinburn, K., Porter, G., & Howard, D. (2014). *Comprehensive Aphasia Test* (L. Haaber Hansen & M. Kaae Frederiksen, adaptation). Copenhagen, Denmark: Dansk Psykologisk Forlag. (Originally published 2005).

Tallberg, I. M. (2005). The Boston naming test in Swedish: Normative data. *Brain and Language*, *94*, 19–31.

Tanaka-Ishii, K., & Terada, H. (2011). Word familiarity and frequency. *Studia Linguistica*, *65*, 96–116.

Tsapkini, K., Vlachou, C. H., & Potagas, C. (2010) Adaptation and validation of standardized aphasia tests in different languages: Lessons from the Boston diagnostic aphasia examination: Short form in Greek. *Behavioural Neurology*, *22*, 111–119.

Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology*, *67*, 1176–1190.

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

Whitworth, A., Webster, J., & Howard, D. (2014). *A cognitive neuropsychological approach to assessment and intervention in aphasia: A clinician's guide*. Hove, UK: Psychology Press.

Ziegler, J. C., Perry, C., & Coltheart, M. (2000). The DRC model of visual word recognition and reading aloud: An extension to German. *European Journal of Cognitive Psychology*, *12*, 413–430.
Ziegler, J. C., Perry, C., & Coltheart, M. (2003). Speed of lexical and nonlexical processing in French: The case of the regularity effect. *Psychonomic Bulletin & Review*, *10*, 947–953.
Zipf, G. K. (1949). *Human behaviour and the principal of least effort*. Cambridge, MA: Addison-Wesley.