



**UvA-DARE (Digital Academic Repository)**

**Visual and auditory digit-span performance in native and nonnative speakers**

Olsthoorn, N.M.; Andringa, S.J.; Hulstijn, J.H.

*Published in:*  
International Journal of Bilingualism

*DOI:*  
[10.1177/1367006912466314](https://doi.org/10.1177/1367006912466314)

[Link to publication](#)

*Citation for published version (APA):*

Olsthoorn, N. M., Andringa, S., & Hulstijn, J. H. (2014). Visual and auditory digit-span performance in native and nonnative speakers. *International Journal of Bilingualism*, 18(6), 663-673. DOI: 10.1177/1367006912466314

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <http://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

# International Journal of Bilingualism

<http://ijb.sagepub.com/>

---

## **Visual and auditory digit-span performance in native and non-native speakers**

Nomi M. Olsthoorn, Sible Andringa and Jan H. Hulstijn

*International Journal of Bilingualism* 2014 18: 663 originally published online 7 December 2012

DOI: 10.1177/1367006912466314

The online version of this article can be found at:

<http://ijb.sagepub.com/content/18/6/663>

---

Published by:



<http://www.sagepublications.com>

**Additional services and information for *International Journal of Bilingualism* can be found at:**

**Email Alerts:** <http://ijb.sagepub.com/cgi/alerts>

**Subscriptions:** <http://ijb.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

**Citations:** <http://ijb.sagepub.com/content/18/6/663.refs.html>

>> [Version of Record](#) - Nov 18, 2014

[OnlineFirst Version of Record](#) - Dec 7, 2012

[What is This?](#)

# Visual and auditory digit-span performance in native and non-native speakers

International Journal of Bilingualism

2014, Vol. 18(6) 663–673

© The Author(s) 2012

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/1367006912466314

ijb.sagepub.com



**Nomi M. Olsthoorn**

Amsterdam Center for Language and Communication, University of Amsterdam, The Netherlands;

Department of Psychology, Lancaster University, UK

**Sible Andringa and Jan H. Hulstijn**

Amsterdam Center for Language and Communication, University of Amsterdam, The Netherlands

## Abstract

We compared 121 native and 114 non-native speakers of Dutch (with 35 different first languages) on four digit-span tasks, varying modality (visual/auditory) and direction (forward/backward). An interaction was observed between nativeness and modality, such that, while natives performed better than the non-natives on the auditory tasks (which were performed in the non-natives' second language), performance on the visual tasks (which was performed in participants' dominant language) did not significantly differ between natives and non-natives. The interaction between nativeness and modality disappeared when the data were corrected for Dutch proficiency. Correction for Dutch proficiency elevated non-native speakers' scores on the auditory tasks, without altering the non-natives' digit-span rank order. Despite considerable differences in mean length of the digit names zero to nine in the non-natives' first languages, these differences were not significantly correlated with their visual digit-span scores. While further research is needed on the sources of variation in digit-span performance, we recommend the use of the visual digit-span task (forward or backward) for cross-linguistic research and advise researchers to be aware of the association between language proficiency and verbal working-memory performance.

## Keywords

Bilingual digit span, visual digit span, auditory digit span, forward digit span, backward digit span, verbal working memory

---

## Corresponding author:

Jan H. Hulstijn, Amsterdam Center for Language and Communication, University of Amsterdam, Spuistraat 120, 1012VT Amsterdam, The Netherlands.

Email: [j.h.hulstijn@uva.nl](mailto:j.h.hulstijn@uva.nl)

## Introduction

Linguistic research into individual differences in language proficiency faces the challenge of finding the factors responsible for inter-individual variations. One of the factors often mentioned as a covariate is working-memory (WM) capacity, more specifically verbal WM capacity. However, many measures of verbal WM tend to be affected by long-term memory factors such as vocabulary knowledge and familiarity with the speech sounds of a language. This is particularly an issue when participant samples are heterogeneous. This paper addresses the question of whether the visual digit span could be a simple verbal WM measure that can be used reliably with native and non-native speakers alike, testing verbal WM but not at the same time linguistic proficiency.

First introduced in 1974 by Baddeley and Hitch, verbal WM is just one of three modality-based components subserving the Central Executive (for an overview, see Baddeley, 2003). Also called the phonological loop, the verbal WM component allows us to store and retrieve a limited amount of phonological information, for a brief period of time. This time period can be extended by the use of sub-vocal rehearsal. Verbal WM memory capacity has shown itself to be a good predictor of successful foreign-language learning (e.g. Atkins & Baddeley, 1998; Gathercole, Service, Hitch, Adams, & Martin, 1999; Kormos & Safar, 2008; Service, 1992; Service & Kohonen, 1995), as well as linguistic aptness in the first language (e.g. Daneman & Carpenter, 1980; Just & Carpenter, 1992; Waters & Caplan, 1996, 2005; although see also MacDonald & Christiansen, 2002, for an alternative view on this matter).

Verbal WM is also assumed to play a role in linguistic proficiency, both for speakers of a first language (L1) and speakers of a second language (L2); however, measuring verbal WM in a multilingual setting has proven difficult. The problem lies with selecting a task that measures verbal WM, but not simultaneously language proficiency. Unfortunately, most if not all measures of verbal WM by definition utilise stimuli that are linguistic in some way or other, be they actual words (as in word or sentence spans), non-words or digits. Consequently, they tap into long-term memory for linguistic knowledge, in addition to WM per se. This results in an obvious advantage for those familiar with the language in question. Below we will briefly discuss the current measures and the disadvantages of each when used in a multilingual setting.

Sentence-span tests (Daneman & Carpenter, 1980), like most WM tasks, typically are presented in series: two sentences of average length, followed by a series of three such sentences, followed by four, etc. The participant listens to the sentences or reads them aloud and after each series has to recall the final word of each sentence. After having completed three series of a particular length, the participant then moves on to the next length, and so on. Sentence span is determined as the last length of which a participant has correctly recalled all words from at least two out of three series. Word-span tests work on a similar principle but use lists of words, rather than full sentences as the initial input.

The problem with both the sentence-span and the word-span task is that they present an advantage for native speakers that know the language over non-native speakers who might not be familiar with some of the words or the predictability of sentential or morphophonological structures. A solution to this is using non-words instead of existing words. The non-word task has two variants: non-word repetition and non-word recognition (Gathercole, Pickering, Hall, & Peaker, 2001). Both make use of non-words as stimuli, in lists of increasing length that are presented auditorily and subsequently have to be either recalled from memory (for the repetition task), or compared to a second list in which the same items are either in identical order as in the first list, or with two items of the list swapped in position (for the recognition task). In the list-comparison task, participants

respond with a YES (they are the same) or NO (they are not the same), said aloud or indicated via a button press.

The use of non-words as stimuli in a verbal WM test is also problematic. Even though non-words are by definition not real words, and therefore unfamiliar to both native speakers and non-native speakers, non-words in one language may be (part-) words in another. This even applies to artificially generated simple consonant–vowel–consonant (CVC) structures. It is therefore not surprising that familiarity and neighbourhood effects for non-word recall have been widely reported (e.g. Hulme, Maughan, & Brown, 1991; Hulme, Roodenrys, Brown, & Mercer, 1995; Roodenrys, Hulme, & Brown, 1993). Even if the items are not evoking any existing words, non-word recall is also affected by familiarity with the phonotactic properties and frequencies of the language from which the non-words are generated (Gathercole, 1995; Gathercole, Frankish, Pickering, & Peaker, 1999; Gathercole et al., 2001; Gathercole, Willis, Emslie, & Baddeley, 1991; Kovács & Racsomány, 2008; Roodenrys & Hinton, 2002; van Bon & van der Pijl, 1997), giving participants whose L1 is more similar to the non-word source an advantage.

Daneman and Merikle (1996) found that numerical stimuli could also be used reliably as items in a measure of verbal WM. This is the basis of digit-span tasks (e.g. Salthouse, Mitchell, Skovronek, & Babcock, 1989). Again, participants are presented with series of stimuli (digits) of increasing length, which they have to recall. Each length is presented two times, and a participant's span is defined as the last length at which one or both lists from a series of a particular length have been recalled correctly. Variations include visual and auditory presentation of the stimuli, and recall being required in forward or backward order with reference to the original list. The forward span task is thought to reflect mainly storage, and the backward span reflects processing aspects of WM as well.

Research on digit-span performance in bilinguals has shown that familiarity with the language plays an important role in digit-span performance, with participants generally achieving higher span scores in their mother tongue than in their non-dominant language(s) (Chincotta & Hoosain, 1995; Chincotta & Underwood, 1997b; Da Costa Pinto, 1991; Thorn & Gathercole, 2001). This is in spite of the fact that digits can be considered to be relatively overlearned stimuli in any language and are usually among the very first words taught in foreign-language education.

Digit span has been studied in relation to linguistic diversity in several ways. Cross-linguistic comparisons show that digit-span size differs across languages, depending on between-language variation in the length of digit names and thus in the average duration to overtly or covertly articulate digit names. For example, digit names in Chinese consist of fewer phonemes (average of 2.5 phonemes) than digit names in English (3.2), while digit names in English are shorter than those in Finnish (5.7) (e.g. Chincotta & Underwood, 1997a; Hoosain, 1979, 1982, 1984; Naveh-Benjamin & Ayres, 1986). Research typically involves comparing participants' span sizes with and without articulatory suppression. Differences in the span size of bilinguals' L1 and L2 are eliminated when the digit-span task is performed under articulatory suppression (Baddeley, Thomson, & Buchanan, 1975; Chincotta & Underwood, 1997a). Additional research conducted by Chincotta, Hyöna, and Underwood (1997) showed that, if digit names are longer in subjects' L1 than in their L2, which is the case in Finnish–Swedish bilinguals, the advantage of performing a digit-span task in L1 over performing it in L2 is smaller than if digit names are shorter in L1 than in L2, which is the case in Swedish–Finnish bilinguals. This finding suggests that not only actual word length of the digits and subjects' speech rate are the source of variation, as was previously postulated (Ellis & Hennely, 1980), but also subjects' 'relative level of competence between the languages for numerals' (Chincotta et al., 1997, p. 257).

In principle, the research just mentioned poses two potential problems for comparative research on the contributing factors of linguistic skills. Firstly, the between-language articulation-speed effect suggests that it might not be possible to compare speakers of different languages on their digit-span performance, because these languages may vary in word length and in average articulation rate. This challenges the possibility (1) of using a heterogeneous group of L2 learners in linguistic-proficiency research and (2) of comparing a group of L2 speakers (of either homogeneous or heterogeneous language background) of a given language with native speakers of that language. Secondly, comparing native and non-native speakers is also complicated by the finding that digit-span performance is generally better in one's L1 than in one's L2.

Despite the potential problems just mentioned, the present study explores the possibility of using the visual and auditory digit-span task for comparing verbal WM span between speakers of different language backgrounds. Participants were 114 non-native speakers of Dutch (native speakers of a large variety of languages) and 121 native speakers of Dutch. The study reported here is part of a research project aiming at explaining individual differences between and within native and non-native speakers of Dutch in performing a listening comprehension task in Dutch. The explanation of individual differences in listening comprehension is sought in participant characteristics such as age, level of education, non-verbal intelligence and WM capacity, and in performance on a range of verbal tasks (one of which was a Dutch vocabulary-size task) (Andringa, Olsthoorn, van Beuningen, Schoonen, & Hulstijn, 2012). Thus, one important question in the project was which type of WM task we should best use to predict individual differences in listening comprehension. To our knowledge, the present study is the first one investigating differences between visual and auditory digit-span performance in non-native speakers in comparison to native speakers, examining the possible mediating effect of language proficiency (alternatively called language competence or language familiarity). In addition, the study allowed us to replicate an effect of between-language differences in the average length of digit names on visual digit span.

Visual digit-span tasks are widely used as a measure of the phonological loop. The phonological similarity effect for visual stimuli (Baddeley, 1968) and the finding that irrelevant speech sounds tend to disrupt memory even for visually presented digits (Colle & Welch, 1976; Salame & Baddeley, 1982) demonstrate that sub-vocal rehearsal is involved in memorising visually presented items. While the auditory digit-span task forces participants to encode the information in the language of the stimuli, the visual digit-span task allows participants to encode the sequences of digits in their preferred language. This may make the visual task more suitable for measuring verbal WM capacity in a multilingual setting, that is, with participants with different language backgrounds, even though between-language differences in the length of digit names may affect performance. We also wanted to explore whether the advantage of native speakers over non-native speakers in the auditory task might decrease when scores are corrected for language proficiency. The expectation was that correction for language familiarity (i.e. L2 proficiency in the case of the non-native speakers) will not affect the native–non-native comparison when the digit-span task is performed in the visual mode, because participants will perform the task in their dominant language (L1). In contrast, correction for language familiarity might compensate for non-native speakers' disadvantage of performing the task in the auditory mode.

## Method

### Participants

Participants were 121 native speakers of Dutch (84 female) and 114 non-native speakers (76 female). All non-native speakers were post-puberty learners of Dutch as a foreign language and their proficiency level was equivalent to or higher than B1 in terms of the *Common European*

*Framework of Reference for Languages* (Council of Europe, 2001). They were native speakers of 35 different languages, including Russian (9 speakers), German (9 speakers), Bahasa Indonesia (9 speakers) and Spanish (8 speakers). All participants were selected from the same age range (19–40), but the native group was slightly younger ( $M = 25$ ;  $SD = 5$ ) than the non-native group ( $M = 29$ ;  $SD = 5$ ). This difference was not significant, however. Participants of both high (higher vocational and university) and low (all other vocational) educational backgrounds were recruited (for the natives, 61 high and 60 low; for the non-natives, 66 high and 47 low). Despite this, the groups differed on the complex matrices component from the Wechsler Adult Intelligence Scale III (WAIS-III), a measure of non-verbal IQ (Wechsler, 1997), with native speakers scoring slightly higher ( $M = 11.8$ ;  $SD = 2.5$ ) than the non-native speakers ( $M = 10.9$ ;  $SD = 3.0$ ); this difference was significant ( $F(1, 233) = 6.588$ ;  $p < 0.05$ ).

## Materials

*Auditory forward and backward digit-span tasks.* Dutch language digits one to nine were synthesised by using the MBROLA speech synthesiser (Dutoit, Pagel, Pierret, Bataille, & Van der Vrecken, 1996) with the NL3 diphone database, and all were subsequently edited in Praat (Boersma & Weenink, 2009) to be exactly 1 second in duration. All digits were then judged by three independent native speakers on their comprehensibility, and adapted accordingly if deemed necessary until all three judges were in agreement, with the provision that all resulting sound files would remain 1 second in duration.

The procedures for the auditory forward and backward digit-span tasks followed the recommendations of the WAIS-III Digit Span subtasks (Wechsler, 1997) and were implemented using the e-prime experimental software package (Schneider, Eschman, & Zuccolotto, 2002). The only change was that participants were required to type in their responses, rather than say them out loud.<sup>1</sup> Participants were instructed that they would hear series of numbers and that their task was to enter these numbers in the correct order into the computer. Participants were seated at about 60 cm from a computer screen and were wearing headphones. To familiarise the participants with the digitised numbers, these were played in order from one to nine, while at the same time their visual counterpart was displayed on the monitor in large font for 1 second. After having been familiarised with the auditory digits, participants then received three practice trials (consisting of two trials of length 2 and one trial of length 3). In the experiment, participants heard series of digits of increasing length, played back at a rate of one digit per second, and were required to enter these when prompted after each series, in the order in which they were first presented for the forward task and in reverse order for the backward task. The minimum series length was two digits, increasing with one digit every two trials until the maximum length of nine digits (for forward series, eight for backward series) was reached, or until participants failed to respond correctly to both trials of a particular length. Depending on this variable exit point, this part of the experiment took about 5–10 minutes to complete.

*Visual forward and backward digit span.* The procedure for the visual digit-span tasks was exactly the same as for the auditory tasks, but instead of audio files of word names that were being played, the numbers (numerals) were presented in 75-point Arial font in blue on the centre of a white computer screen for a duration of 1 second each, with no blank screen in between the items of a list. Visual span sessions also omitted the familiarisation block preceding the practice trials and again took between 5 and 10 minutes to complete. Participants were instructed to type in the series of numbers



(in the order required). It was suggested that they use their L1 to perform the task: 'You may do the task in your first language' (literal translation of the original Dutch instruction).<sup>2</sup>

In agreement with the WAIS-III instructions for digit-span tests, WM capacity was determined as the maximum length at which a participant had responded at least one out of two trials of the same length correctly, before having made an incorrect response on both consecutive trials of the same length. Thus, WM capacity could vary from 2 to 9 for the forward tasks, and from 2 to 8 for the backward tasks.

**Vocabulary test.** Vocabulary knowledge was tested with a computer-administered multiple-choice test. For each item, participants had five alternative paraphrases of the meaning of the target word to choose from, the last one always being: 'I really don't know'. The paraphrases were always couched in simple, common language. Several considerations guided item selection and construction. Firstly, words should not be too domain-specific. Care was taken not to introduce any regularity in the length of the alternatives and the way meanings were paraphrased. The final test consisted of 60 items selected from a total of 140 items that were piloted amongst 48 native and non-native speakers. The target words were selected on the basis of frequency information from the CELEX database (Baayen, Piepenbrock, & Gulikers, 1995). CELEX consists of 124,000 lemmas of Dutch, with frequency information based on a corpus of written Dutch of 50-million word forms (see also <http://celex.mpi.nl/>).

## Results

All tests reported below are two-tailed. Scores on the vocabulary test ranged from 26 to 60 for native speakers ( $M = 41$ ;  $SD = 6$ ) and from 4 to 55 for non-native speakers ( $M = 28$ ;  $SD = 9$ ) (*Welch's F* (1, 194.69) = 145.802;  $p < 0.001$ ).

An analysis of variance (ANOVA) of the digit-span data, with direction (2) and modality (2) as within-subject factors, and nativeness (2) as a between-subject factor, revealed a between-subjects main effect of nativeness, with native speakers outperforming the non-native speakers on average by .348 items ( $F$  (1, 233) = 7.09;  $p < 0.05$ ). A within-subjects main effect of direction ( $F$  (1, 233) = 165.668;  $p < 0.001$ ;  $\eta^2_p = .416$ ) also obtained. This effect was due to the forward spans being on average 0.829 items longer than the backward spans. In addition to this, there was an interaction effect between modality and nativeness ( $F$  (1, 233) = 10.275;  $p < 0.01$ ;  $\eta^2_p = .042$ ). No other main or interaction effects were significant. Additional ANOVAs showed that the modality  $\times$  nativeness interaction effect was due to the native speakers and non-native speakers differing on the auditory digit spans, both in the forward direction ( $F$  (1, 233) = 12.76;  $p < 0.0001$ ;  $\eta^2_p = 0.52$ ) and backwards ( $F$  (1, 233) = 11.332;  $p < 0.01$ ;  $\eta^2_p = 0.46$ ) but not on the visual span tasks. Also, native speakers performed better on the auditory tasks ( $F$  (1, 120) = 17.77;  $p < 0.0001$ ;  $\eta^2_p = 0.129$ ), but native and non-native speakers performed the same on the auditory and the visual tasks.

Since the native and non-native groups showed an a priori difference in IQ, and IQ and vocabulary were found to be correlated ( $r = .342$ ;  $p < 0.001$ ), the IQ component of vocabulary was partialled out by means of linear regression analysis, and the residuals were saved. To check for effects of familiarity with the language, another factorial analysis was conducted with the vocabulary residuals as a covariate (see Table 1 for both original means and least-square means). This resulted in a main effect of direction ( $F$  (1, 232) = 165.189;  $p < 0.01$ ;  $\eta^2_p = 0.416$ ), but no other main or interaction effects approached significance. In all eight measures (combinations of three binary distinctions: natives/non-natives, auditory/visual, and forward/backward), the correlations between



the raw scores and the scores corrected for language proficiency (Dutch vocabulary) were in the range of Pearson  $r$  values between .96 and 1.00. Thus, while correction for language proficiency raised non-natives' scores in the auditory tasks (see Table 1), subjects' rank order remained virtually the same.

The associations between the four original digit-span measures were moderate in both the native and the non-native groups, with Pearson  $r$  values ranging between .45 and .51 (natives) and between .34 and .61 (non-natives), with all coefficients significant at the .000 level.

The estimated mean length (in phonemes) of the 10 digit names (from zero to nine) in participants' L1 varied from 2.5 (Mandarin) to 6.0 (Kinyarwanda). The mean length calculated over all 114 non-native participants was 3.9 ( $SD = 0.6$ ). The mean phoneme length of the 10 digit names in Dutch is 3.1. Importantly, mean phoneme length and visual digit span were not significantly associated in the non-natives ( $n = 114$ ); Pearson  $r$  values between mean phoneme length and forward and backward digit span was  $-.06$  and  $-.16$  ( $ns$ ).

## Discussion

The main aim of the study was to find out which type of digit-span task lends itself best for use in studies involving both native and non-native speakers of a given language. Adult native and non-native speakers of Dutch ( $n = 121$  and  $114$ , respectively) performed four digit-span tasks (auditory forward, auditory backward, visual forward and visual backward). As expected, a significant main effect of direction was obtained. Furthermore, a significant nativeness  $\times$  modality effect was observed: while natives obtained significantly higher scores than non-natives in the auditory mode, this was not the case in the visual mode. This significant interaction (i.e. the advantage of the natives in the auditory mode) disappeared when performance was corrected for language proficiency (the Dutch vocabulary test). Between-language differences in the average length of digit names in non-native participants' L1 (ranging between 2.5 and 6.0 phonemes) did not appear to affect their performance in the two visual digit-span tasks.

The results confirm a well-documented effect of direction, with forward tasks generally resulting in longer spans than backward tasks (see Ramsay & Reynolds, 1995, for a review). We also replicated the effect of proficiency, that is, that native speakers tend to outperform non-native speakers in digit-span tasks (Chincotta & Underwood, 1997b; Da Costa Pinto, 1991; Thorn & Gathercole, 2001). In our study, the advantage only held for the auditory version of the task because that task had to be performed in Dutch, while the visual task could be performed in participants' native language. On average, digit span in the visual modality was equal for the native and non-native participants, despite the fact that the mean digit-name length of the native

**Table 1.** Average span scores (SD in brackets) for native and non-native speakers and least-square means for the groups corrected for vocabulary score with IQ partialled out.

|                   | <i>ns</i>    | <i>nns</i>   | <i>ns</i> least-square means | <i>nns</i> least-square means |
|-------------------|--------------|--------------|------------------------------|-------------------------------|
| Forward auditory  | 6.45 (1.118) | 5.91 (1.209) | 6.415 (.120)                 | 5.954 (.124)                  |
| Forward visual    | 6.07 (1.243) | 5.89 (1.431) | 6.032 (.138)                 | 5.922 (.143)                  |
| Backward auditory | 5.55 (1.204) | 4.98 (1.395) | 5.465 (.133)                 | 5.076 (.138)                  |
| Backward visual   | 5.28 (1.260) | 5.18 (1.473) | 5.331 (.141)                 | 5.131 (.146)                  |

languages of the non-native participants was higher (3.9) than the mean digit-name length in the Dutch language (3.1).

The fact that natives obtained higher scores in the auditory than in the visual tasks is reflective of an auditory superiority effect (Penney, 1989). The auditory superiority effect for serial recall items is generally thought to reflect the fact that auditory presentation leads more reliably to the creation of a vivid phonological code than does visual presentation of numerals, because subjects have to access and retrieve the word names in the visual task, while the word names are presented in the auditory task (Baddeley, 1986; Gathercole & Baddeley, 1990). The auditory superiority effect was not obtained with the non-native speakers. Apparently, the advantage of being presented with the word names in the auditory task was cancelled because the task was performed in their L2.

The fact that all differences between native and non-native speakers disappeared when linguistic proficiency was taken into account is telling. It suggests that even with such overlearned and abstract stimuli as digits, the phonological-loop capacity, measured with a digit-span task, reflects, at least partly, familiarity with (proficiency in) the language. As such, our results are corroborative of MacDonald and Christiansen's (2002) claim that verbal WM equals one's amount of experience with comprehension processes in a particular language. The following important question then arises. Is it empirically possible to tease apart, in digit-span performance in a group of native speakers, (i) the variance produced by individual differences in articulation speed and (ii) the variance produced by individual differences in familiarity with the language? In other words, to what extent is a person's articulation speed an index of his or her language familiarity?

On the basis of the literature and on the basis of our findings, what advice can we give to researchers of bilingualism with respect to measuring verbal WM capacity in a cross-linguistic setting? In the introduction, we listed some important disadvantages of using sentence-span tasks and non-word-span tasks. In general, depending on the study's research questions of course, the auditory and visual digit-span tasks appear to lend themselves better to the purpose. The advantage of the auditory digit-span task is that all participants conduct the task in the same language. Researchers might consider correcting the raw digit-span scores with a measure of language proficiency (such as vocabulary). Importantly, in our study this did not affect participants' digit-span rank order. The advantage of the visual digit-span task (with numerals as stimuli) is that participants perform the task in their strongest language, and thus the score comes closer to the success with which they can store and process phonological information in WM. There is strong evidence that this type of WM capacity is positively associated to learning a foreign language (see the references in the introduction). The potential drawback of the fact that languages differ in the length of numeral names can be dealt with by estimating the mean length of digit names in participants' first languages and computing the association between that length and participants' visual digit-span score. Depending on the study's cross-linguistic participant sample, this may or may not form an obstacle. Surprisingly, in our study, the visual digit span of the non-natives was not significantly different from that of the natives, although the average digit-name length in their first languages was longer than that in Dutch. These findings may encourage researchers to use the visual digit-span tasks despite between-language differences in digit-name length. The decision to use the forward or backward version of the task will depend on the researcher's wish to increase (backward) or not (forward) the WM component of information manipulation.

### **Acknowledgements**

The authors wish to thank Netta Meijer and Tineke van der Linde for their help running the experiments and Padraic Monaghan for his help in generating the auditory digits.

## Funding

This work was supported by the Dutch Organisation for Scientific Research (NWO; grant number 360-70-230).

## Notes

1. As one anonymous reviewer of this paper noted, having the participants type their responses is a non-standard implementation of the verbal digit-span task, because of the recoding of verbal information into motor and spatial codes necessary for the response. Also, due to the requirement to give typed responses, the participants may have been prone to encode verbal stimuli into a series of spatial positions before storage and thus make use of the Visual-Spatial Sketchpad component of WM.
2. Unfortunately we failed to ask the non-native participants afterwards whether they had performed the visual digit-span tasks in their dominant language.

## References

- Andringa, S., Olsthoorn, N., Van Beuningen, C., Schoonen, R., & Hulstijn, J. (2012). Determinants of success in native and non-native listening comprehension: An individual differences approach. *Language Learning*, 62 [Supplement], 49–78.
- Atkins, P. W. B., & Baddeley, A. D. (1998). Working memory and distributed vocabulary learning. *Applied Psycholinguistics*, 19, 537–552.
- Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database (CD-ROM)*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.
- Baddeley, A. D. (1968). How does acoustic similarity influence short-term memory? *Quarterly Journal of Experimental Psychology*, 20, 249–264.
- Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.
- Baddeley, A. D. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, 36, 189–208.
- Baddeley, A. D., & Hitch, G. J. (1974). Working Memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47–89). New York, NY: Academic Press.
- Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of working memory. *Journal of Verbal Learning and Verbal Behavior*, 14, 575–589.
- Boersma, P., & Weenink, D. (2009). *Praat: doing phonetics by computer* (Version 5.1.20) [Computer program]. Retrieved from <http://www.praat.org/>
- Chincotta, D., & Hoosain, R. (1995). Reading rate, articulatory suppression and bilingual digit span. *European Journal of Cognitive Psychology*, 7, 201–211.
- Chincotta, D., Hyöna, J., & Underwood, G. (1997). Eye fixations, speech rate and bilingual digit span: Numeral reading indexes fluency not word length. *Acta Psychologica*, 9, 253–275.
- Chincotta, D., & Underwood, G. (1997a). Digit span and articulatory suppression: A cross-linguistic comparison. *European Journal of Cognitive Psychology*, 9, 89–96.
- Chincotta, D., & Underwood, G. (1997b). Speech rate estimates, language of schooling and bilingual digit span. *European Journal of Cognitive Psychology*, 9, 325–348.
- Colle, H. A., & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning and Verbal Behavior*, 15, 17–32.
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge, UK: Cambridge University Press.
- Da Costa Pinto, A. (1991). Reading rates and digit spans in bilinguals: The superiority of the mother tongue. *International Journal of Psychology*, 26, 471–483.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450–466.
- Daneman, M., & Merikle, P. (1996). Working memory and language comprehension: A meta-analysis. *Psychological Bulletin & Review*, 3, 422–433.

- Dutoit, T., Pagel, V., Pierret, N., Bataille, F., & Van der Vrecken, O. (1996). The MBROLA project: Towards a set of high quality speech synthesizers free of use for non commercial purposes. In H. T. Bunell & W. Isardi (Eds.), *Proceedings of the Fourth International Conference on Spoken Language Processing* (pp. 1393–1396).
- Ellis, N. C., & Hennely, R. A. (1980). A bilingual word-length effect: Implications for intelligence testing and the relative ease of mental calculation in Welsh and English. *British Journal of Psychology*, *71*, 43–51.
- Gathercole, S. E. (1995). Is nonword repetition a test of phonological memory or long-term knowledge? It all depends on the nonwords. *Memory & Cognition*, *23*, 83–94.
- Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language*, *29*, 336–360.
- Gathercole, S. E., Frankish, C. R., Pickering, S. J., & Peaker, S. (1999). Phonotactic influences on short-term memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *25*, 84–95.
- Gathercole, S. E., Pickering, S. J., Hall, M., & Peaker, S. (2001). Dissociable lexical and phonological influences on serial recognition and serial recall. *Quarterly Journal of Experimental Psychology*, *54A*, 1–30.
- Gathercole, S. E., Service, E., Hitch, G. J., Adams, A. M., & Martin, A. J. (1999). Phonological short-term memory and vocabulary development: Further evidence on the nature of the relationship. *Applied Cognitive Psychology*, *13*, 65–77.
- Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1991). The influence of syllables and word-likeness on children's repetition of nonwords. *Applied Psycholinguistics*, *12*, 349–367.
- Hoosain, R. (1979). Forward and backward digit span in the languages of the bilingual. *Journal of Genetic Psychology*, *135*, 263–268.
- Hoosain, R. (1982). Correlation between pronunciation speed and digit span size. *Perceptual & Motor Skills*, *55*, 1128.
- Hoosain, R. (1984). Experiments on digit spans in the Chinese and English languages. In H. S. R. Kao & R. Hoosain (Eds.), *Psychological Studies of the Chinese language* (pp. 23–28). Hong Kong: The Chinese Language Society of Hong Kong.
- Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, *30*, 685–701.
- Hulme, C., Roodenrys, S., Brown, G., & Mercer, R. (1995). The role of long-term memory mechanisms in memory span. *British Journal of Psychology*, *86*, 527–536.
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension. *Psychological Review*, *99*, 122–149.
- Kormos, J., & Safar, A. (2008). Phonological short term memory, working memory and foreign language performance in intensive language training. *Bilingualism: Language and Cognition*, *11*, 261–271.
- Kovács, G., & Racsmany, M. (2008). Handling L2 input in phonological STM: The effect of non-L1 phonetic segments and non-L1 phonotactics on nonword repetition. *Language Learning*, *58*, 597–624.
- MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*, *109*, 35–54.
- Naveh-Benjamin, M., & Ayres, T. J. (1986). Digit span, reading rate and linguistic relativity. *Quarterly Journal of Experimental Psychology*, *38*, 739–751.
- Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory and Cognition*, *17*, 398–422.
- Ramsay, M. C., & Reynolds, C. R. (1995). Separate digits tests: A brief history, a literature review and a reexamination of the factor structure of the test of memory and learning (TOMAL). *Neuropsychology Review*, *5*, 151–171.
- Roodenrys, S., & Hinton, M. (2002). Sublexical or lexical effects on serial recall of nonwords? *Journal of Experimental Psychology: Learning, Memory and Cognition*, *28*, 29–33.
- Roodenrys, S., Hulme, C., & Brown, G. (1993). The development of short-term memory span: Separable effects of speech rate and long-term memory. *Journal of Experimental Child Psychology*, *56*, 431–442.

- Salame, A. P., & Baddeley, A.D. (1982). Disruption of short-term memory by irrelevant speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior*, 21, 150–164.
- Salthouse, T. A., Mitchell, D. R. D., Skovronek, E., & Babcock, R. L. (1989). Effects of adult age and working memory on reasoning and spatial abilities. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15, 507–516.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-prime reference guide*. Pittsburgh, PA: Psychology Software Tools Inc.
- Service, E. (1992). Phonology, working memory and foreign-language learning. *Quarterly Journal of Experimental Psychology*, 44A, 21–50.
- Service, E., & Kohonen, V. (1995). Is the relation between phonological memory and foreign language learning accounted for by vocabulary acquisition? *Applied Psycholinguistics*, 16, 155–172.
- Thorn, A. S. C., & Gathercole, S. E. (2001). Language differences in verbal short-term memory do not exclusively originate in the process of subvocal rehearsal. *Psychonomic Bulletin & Review*, 8, 357–364.
- van Bon, W. H. J., & van der Pijl, J. M. L. (1997). Effects of word length and wordlikeness on pseudoword repetition by poor and normal readers. *Applied Psycholinguistics*, 18, 101–114.
- Waters, G., & Caplan, D. (1996). The capacity theory of sentence comprehension: Critique of Just and Carpenter (1992). *Psychological Review*, 4, 761–772.
- Waters, G., & Caplan, D. (2005). The relationship between age, processing speed, working memory capacity and language comprehension. *Memory*, 13, 403–413.
- Wechsler, D. (1997). *WAIS-III administration and scoring manual*. San Antonio, TX: The Psychological Corporation.

### Author biographies

Nomi M. Olsthoorn has a background in Cognitive Psychology. Her PhD research concerned the relationship between language production and perception and she has since focussed on the cognitive correlates of second language processing. She currently teaches in psychology and linguistics at Lancaster University.

Sible Andringa is Assistant Professor Dutch Linguistics at the University of Amsterdam, with a special interest in the acquisition of Dutch as a second language. His research focuses on the development of second language proficiency and the role of instruction in proficiency development. Webpage: <http://home.mederwerker.uva.nl/s.j.andringa/>

Jan H. Hulstijn is Professor of Second Language Acquisition at the University of Amsterdam. He is currently writing a book entitled 'Language proficiency in native and non-native speakers: Theory, research, and measurement'. Webpage: <http://home.mederwerker.uva.nl/j.h.hulstijn/>