# Working Memory is more than what is measured by a WM task: a case in point: segmentation proficiency and verbal working memory

Olsthoorn, N.M.; Andringa, S.; Meijer, N.

**Publication date**
2009

**Document Version**
Final published version

**Published in**
Artikelen van de zesde Anéla-conferentie

# Working Memory is more than what is measured by a WM task: A case in point: Segmentation proficiency and verbal working memory

Nomi Olsthoorn, Sible Andringa, Netta Meijer,Universiteit van Amsterdam, ACLC- TTV, Spuistraat 134, 1012 VB Amsterdam, n.m.olsthoorn@uva.nl

## Introduction

Everybody uses language. However, people differ in the extent to which they are proficient at using language. This applies to native speakers as well as second language learners. Measuring these differences is an important task for applied linguistics, but just assessing the differences is not enough, we also want to explain the source of the variation in language proficiency, and for instance, use that knowledge to accurately predict whether a person will do well in learning a second language or not.

One of the variables that is known to correlate with individual differences in language proficiency and with the ability to learn a second language and become proficient in it is one's working memory span, more precisely verbal working memory.

Verbal working memory however, although rather straightforwardly described by Baddeley and others as the part of working memory that for a limited period of time can store a limited amount of auditory information, is far from a unitary concept. The extent to which verbal working memory is related to language proficiency very much depends on both the paradigm and the stimuli chosen to assess the working memory span and selection of the task to measure linguistic proficiency.

In this talk we'll provide an overview of aspects to consider when selecting the appropriate working memory task given a certain linguistic skill. To illustrate, we'll use an example out of our research into listening comprehension, namely a study of segmentation skills as a case in point, and we'll conclude with a reflection on the validity of the status of the verbal working memory as a unitary process.

## Measuring segmentation proficiency

When perceiving auditory input, listeners are subjected to a continuous speech signal. From that signal they have to retrieve word boundaries in order to extract discrete meaningful units, i.e. words. Furthermore, in casual speech, segments are oftentimes reduced sometimes to an extreme extent – which leads to large variations in the realization of speech.

Speech segmentation in general has been frequently studied, in both native (e.g. Cutler 2002; Cutler & Clifton 1999; Nakatani & Schaffer 1978; Van der Lugt 1999) and nonnative listening (e.g. Altenberg 2005; Carroll 2004). Reduction has been shown to delay comprehension in native speakers of Dutch (Ernestus & Baayen 2007; Ernestus, Baayen & Schreuder 2002), but so far, the effects in non-natives have not been investigated systematically. We studied speech segmentation in both native and non-native listening. The research question was whether there is a difference between native and non-native speakers of Dutch in the segmentation of reduced and unreduced Dutch forms. Two segmentation tasks were constructed; participants were presented with full and reduced auditory input (sentence fragments) and had to indicate how many (by pressing the corresponding number key), as well as which words (by typing them out) they heard. In this second task, speed as well as accuracy of segmentation were measured. Both tasks made use of the same stimulus set. A



typical stimulus fragment was for instance: *ik kan volgende keer*, recorded in natural speech by a young female native speaker of Dutch.

Native and non-native speakers of Dutch took part in the experiment, in which also level of education was varied (in natives and non-natives) as was age (in natives). A pilot study showed that native speakers are faster and more accurate than non-native speakers in segmenting unreduced and reduced speech (Meijer 2008). This may be ascribed to the inability of non-natives to effectively use Dutch segmentation cues. They may also be less able to use reconstructive processes, as is indicated by a larger effect of reduction than in the native group. A notable finding was the interaction between nativeness and task type. Non-native speakers were better at counting words than at identifying them. This may represent a phase in the what seems to be a gradual process in the acquisition of segmentation skills. Analysis of the three reduction types showed that extreme reductions are the most difficult to segment. There was no difference between vowel and consonant reduction.

However, finding that native speakers outperform non-natives on tasks like these is not surprising. Therefore, apart from measuring how well native speakers and non-native speakers do on these segmentation tasks, we also want to see whether good performance somehow correlates with working memory span. Introspectively, when doing our segmentation task – even in an informal setting – listening to a sentence fragment and then either typing in what you have heard, or indicating how many words were discernable from the input stream are processes that draw heavily on the phonological loop of our working memory. To demonstrate this effectively though, the way in which we measure the working memory capacity of our participants crucially determines the effect. Below we discuss the factors that have to be taken into consideration before choosing a working memory task.

## Verbal Working Memory and Language Comprehension

Like in language proficiency, individuals differ in their working memory capacity, and furthermore, working memory capacity declines with age (see Salthouse 1991 for a review). In language comprehension, as in all other cognitive tasks, working memory plays an important role. According to Waters and Caplan (2005): "Regardless of whether language is written or spoken, the input to the comprehender becomes available over time and temporarily discontinuous parts of the input must be related to one another for language to be understood". Indeed, working memory is involved in such diverse listening comprehension tasks as computing the referent of personal pronouns (Just & Carpenter 1980), inferencing (Mitchell, 1982) and resolution of inconsistencies (Daneman & Carpenter 1983), and may therefore be an important key to understanding individual differences in language proficiency.

However, working memory is by no means a simple concept. In one dominant view it is thought of as consisting of three (sometimes four) components: the Central Executive, subserved by the Visuo-spatial sketchpad, the Phonological loop, and (more recently) the Episodic Buffer (Baddeley & Hitch 1974; Baddeley & Logie 1999; Baddeley 2000). In order to measure the effect of individual differences in working memory on language comprehension one has to find a task that taps into the relevant kind of working memory. In Baddeley's multiple-component model outlined above, the task in question should tap into the phonological loop, the part of working memory that holds verbal-acoustic information. Other researchers argue about the specificity of the working memory resource that serves language comprehension. For instance, Just and Carpenter (1992) find that a single verbal resource underlies all processes involved in language comprehension. On the other hand, Waters and Caplan (1996) identify separate language processing resources: one for "automatic" processes such as constructing syntactic representations and thematic role assignment, and one for verbally mediated tasks.

In deciding which task(s) to use in a study of the cognitive and linguistic determinants of first and second language listening proficiency, five issues about the nature of working memory and working memory research have to be taken into careful consideration. But before we go into those issues, a remark should be made about the selection of appropriate working memory tasks in general. With correlational research into the factors that contribute to language proficiency, ideally one is looking for a task that measures just working memory capacity, and nothing else. Possible confounds are grammatical proficiency, world knowledge, vocabulary knowledge to name but a few. In order for language-external cognitive processes such as working memory to have any explicatory force with respect to individual variation in linguistic capacities, it is essential to measure those processes, and not indirectly, language-internal factors. The aspects mentioned above are all known to affect working memory span in some way or another, and are especially dangerous when comparisons are made between L1 and L2 speakers (see also the section about linguistic specificity, below). It is this maxim that underlies all considerations mentioned here.

*I. Capacity vs. Load studies*

There are two approaches to verbal working memory in language comprehension: memory load studies and memory span studies (Caplan & Waters 1999; Vos, Gunter, Schriefers & Friederici 2001). Working memory load studies investigate the potential effect of a concurrent memory load on language comprehension whereas working memory span studies focus on the potential effect of individual differences in working memory capacity on language comprehension. A typical working memory load task requires a participant to perform a (reading or listening) task while at the same time monitor the stimulus material for one (low load) or more (higher load) words to occur.

Concurrent load tasks can be useful to isolate cognitive slowing due to aging effects, regardless of individual working memory capacity. This can be achieved by letting participants perform the (comprehension) task while at the same time attending to a concurrent task, set at a similarly taxing level for all participants (determined by individual assessment). However, for our purpose, working memory measures that assess individual differences are the natural choice.

*II. Storage vs. Processing*

In Baddeley's model, the phonological loop is fractionated into storage and processing components (Baddeley 2003). The phonological store itself is a very transient storage system that can hold memory traces only temporarily, over a matter of seconds, during which they decay, unless actively refreshed by the processing component, a subvocal rehearsal system. Tasks that measure both storage and processing are reading span and listening span. Storage proper can be more accurately measured by (forward) (non)word and digit span tasks. The latter tasks typically consist of lists of items (be they nonwords, words or single numerals, presented either visually or auditorily), which then have to be recalled in the order in which they were presented. However, there are also variants of these tasks that require participants to recognise instead of recall the strings of items. (The distinction between recognition and recall tasks is a very relevant issue, which will be discussed further in section V, below).

Conversely, combined storage and processing tasks require participants to read or listen to a set of unrelated sentences, and then, at the end of the set, to recall the final word of each sentence in the set. As participants are presented with increasingly longer sets of sentences, working memory capacity is typically defined as the maximum number of sentences the participant can process while maintaining perfect recall of the final words (Daneman & Carpenter 1980). In the literature, this task occurs with or without a comprehension check following each sentence (e.g. Baddeley, Logie, Nimmo-Smith & Brereton 1985). The comprehension check task allows for the stimulus materials to be more or less true (demanding true/false judgements, eg. *Mammals are vertebrates that give birth to live young*), more or less sensible (demanding sensibility judgements, e.g. *The girl sang the apple*) or more or less grammatical (grammaticality judgements).



Individual differences in reading or listening span are theorised to result from the fact that poor comprehenders allocate so much capacity to comprehending the sentences that they can only store two or three final words, whereas good comprehenders typically exhibit a span of around five words.

According to Daneman and colleagues the best predictors for language comprehension are measures that tap into the combined processing and storage capacity of working memory (Daneman & Carpenter 1980; Daneman & Merikle 1996). In their 1996 meta-analysis Daneman and Merikle furthermore showed that the predictive power of storage and processing measures is not limited to tasks that involve only the manipulation of linguistic material, but can also be assessed via mathematical processes as long as the processes involve the manipulation of symbolic information (e.g. words or digits). However, in their analysis, the best predictive validity was obtained by tasks that include a verbal processing component and a verbal storage component. Daneman & Carpenter (1980) thus describe individual differences to occur as a result of a variation in their *functional* capacity of allocating the limited resources that must be shared between storage and processing demands of the (working memory) task at hand.

*III. Language Specificity*

It is assumed that in contrast to the processing component, the storage component of the phonological loop is relatively language-independent (Baddeley 2003; Gathercole 1995; Gathercole 2006). Three tasks that use non-words are the non-word repetition task and non-word serial recall/recognition task. These tasks are measures of storage capacity. The difference between non-word serial recall and recognition is that the former displays a lexicality effect (in that words are generally recalled more easily than non-words), whereas the latter does not display such an effect (recognition is similar for words and non-words (Baddeley 2003; Gathercole, Pickering, Hall & Peaker 2001)). In line with the general requirement mentioned above, it is crucial to select stimuli that do not secretly test linguistic abilities, specific to the language in question, but stimuli that are equally difficult (or easy) to process for both native speakers and non-natives.

*IV. Modality*

One way to solve language specific effects of working memory is to use a non-linguistic storage (and/or storage and processing) task that allows participants to encode the materials in their own preferred language. Non-linguistic working memory tasks can be valid predictors of language comprehension, as long as they involve symbol manipulation (Daneman & Merikle 1996). Of course, this is only possible if presentation is also language-independent, as in a visually presented digit span or mathematical storage and processing task (e.g. Salthouse, Mitchell, Skovronek & Babcock 1989).

Certainly, even visually presented materials call for phonological processing, as demonstrated by the fact that visually presented letters and digits tend to be confused on the basis of their *phonological similarity* rather than their visual similarity (i.e. S and F are confused more often than F and E; Conrad 1964). However, it should be noticed that according to Baddeley (2003) the translation of visual information to a phonological code also involves the subvocal rehearsal process. In other words, a (simple) digit span task that normally only taps into phonological storage now also draws on phonological processing.

*V. Characteristics of the linguistic task at hand.*
The final consideration concerns the non-linguistic characteristics of the linguistic task at hand. For instance, are participants presented with visual or auditive information in the linguistic task? If the linguistic task is visually presented, then it makes sense to select a working memory task that also requires visual processing, but does not make use of any linguistic stimuli, in so far as these may produce an advantage for native speakers, according to consideration III (language specificity), above. The same applies to other task-related characteristics such as the distinctions between recall and recognition memory. It is generally acknowledged in memory literature that being able to recall a set of stimuli generally requires more intensive and deeper processing than being able to recognise whether a stimulus (or string of stimuli) has been presented before (Searleman & Herrmann 1994). If the linguistic task in question is therefore demanding, by requiring participants to actively retrieve the stimuli from memory, a working memory task that merely draws upon recognition is not a valid comparison.

## Implications for the segmentation study

To illustrate the effect of the different considerations, we will present the data of six different working memory measures and how they correlate with our segmentation tasks. The tasks are: non-word recognition with phonetically distinct and phonetically similar simple CVC strings, forward and backward recall digit span, both visually and auditorily presented. We predict that recognition tasks are less predictive than recall tasks. With respect to the nature of the stimulus materials however, both linguistic specificity (or lack thereof) will play a role (comparing digits with non-words), as will the mode of presentation, since visually presented digits offer the benefit of encoding freedom in one's preferred language. This as opposed to auditorily presented stimuli, even overlearned stimuli such as digits, which are more beneficial to native speakers than to non-natives.

## Conclusion

When looking for the factors that contribute to individual differences in language proficiency, working memory capacity is a likely candidate and important contributor to explain variance. However, choosing the right instrument to measure working memory span is dependent on many considerations, both with respect to the stimuli, the desired effects, the linguistic task at hand and participant groups to be compared. In the talk we will show that different working memory measures produce different effects as far as correlation with our segmentation task is concerned, we will explain these variations in correlation in terms of the considerations outlined above and reflect on the thesis that even a seemingly straightforward concept such as working memory is more than whatever is measured by a working memory task.

## References


Altenberg, E. P. (2005), The perception of word boundaries in a second language. *Second Language Research,* 21(4), 325-358.

Baddeley, A.D. (2000), The episodic buffer. A new component of working memory? *Trends in cognitive sciences, 4 (11).* 417-423.

Baddeley, A.D. (2003), Working memory and language: an overview. *Journal of communication disorders*, 36: 189-208.





Baddeley, A.D. & G.J. Hitch (1974), Working Memory, in G.H. Bower (ed.), *The psychology of learning and motivation: advances in research and theory* , Vol. 8: 47-89, New York: Academic Press.

Baddeley, A.D. & R.H. Logie (1999), Working Memory: the Multiple-Component Model, in A. Miyake & P. Shah (eds.), Models of working memory: Mechanisms of active maintenance and control: 28-61, Cambridge: University Press.

Baddeley, A.D., Logie, R.H., Nimmo-Smith, I. & N. Brereton (1985), Components of fluent reading, *Journal of Memory and Language, 24,* 119-131.

Caplan, G. & Waters, D. (1999), Verbal working memory and sentence comprehension. *Behavioral and brain sciences* 22: 77-126.

Carroll, S. E. (2004), Segmentation: learning how to "hear words" in the L2 speech stream, *Transactions of the Philological Society, 102*(2): 227-254.

Conrad, R. (1964), Acoustic confusions in immediate memory, *British Journal of Psychology, 55*: 75-84.

Cutler, A. (2002), Native listeners, *European Review, 10*(1): 27-41.

Cutler, A., & C. Clifton Jr. (1999), Comprehending spoken language: a blueprint of the listener, in C. M. Brown & P. Hagoort (eds.), *The neurocognition of language*, Oxford: Oxford University Press.

Daneman M. & P.A. Carpenter (1980), Individual differences in working memory and reading, *Journal of verbal learning and verbal behavior,* 19: 450-466.

Daneman, M. & P.A. Carpenter (1983), Individual differences in integrating information between and within sentences, *Journal of experimental psychology: learning memory and cognition*, 9: 561-584.

Daneman, M. & P. Merikle (1996), Working memory and language comprehension: a meta-analysis, *Psychological Bulletin & Review*, 3 (4): 422-433.

Ernestus, M., & R.H. Baayen (2007), *The comprehension of acoustically reduced morphologically complex words: the roles of deletion, duration, and frequency of occurrence,* Paper presented at the 16th International Congress of Phonetic Sciences, Saarbrücken, Germany.

Ernestus, M., Baayen, R. H. & R. Schreuder (2002), The recognition of reduced word forms, *Brain and Language*, 81, 162-173.

Gathercole, S.E. (1995), Is non-word repetition a test of phonological memory or long-term knowledge? It all depends on the non-words, *Memory and Cognition,* 23: 83-94.

Gathercole, S.E. (2006), Non-word repetition and learning: the nature of the relationship, *Applied Psycholinguistics* 27: 513-543.

Gathercole, S.E., Pickering, S.J., Hall, M. & S.M. Peaker (2001), Dissociable lexical and phonological influences on serial recognition and serial recall, *Quarterly Journal of Experimental Psychology*, 54: 1-30.

Just, M.A. & P.A. Carpenter (1980), A theory of reading: from eye fixations to comprehension, *Psychological Review,* 99: 122-149.

Meijer, N. (2008), *Segmentation of reduced speech by native and non-native listeners of Dutch*, Master's Thesis, University of Amsterdam.

Mitchell, D.C. (1982), *The process of reading: a cognitive analysis of fluent reading and learning to read*, New York: Wiley.

Miyake, A. & P. Shah (eds.), (1999), *Models of working memory: mechanisms of active maintenance and executive control.* Cambridge: University Press.

Nakatani, L. H. & J.A. Schaffer (1978), Hearing "words" without words: prosodic cues for word perception, *The Journal of the Acoustical Society of America, 63*(1): 234-245.

Salthouse, T.A. (1991), Mediation of adult age differences in cognition by reductions in working memory and speed of processing, *Psychological Science,* 2: 179-183.

Salthouse, T.A., Mitchell, D.R.D., Skovronek, E. & R.L. Babcock (1989), Effects of adult age and working memory on reasoning and spatial abilities*, Journal of Experimental Psychology: learning, memory and Cognition,* 15: 507-516.

Searleman, A. & D. Herrmann (1994), *Memory from a broader perspective*, New York: McGraw-Hill.

Van der Lugt, A. (1999), *From speech to words*, Nijmegen: Doctoral dissertation, Katholieke Universiteit Nijmegen.

Vos, S.H., Gunter, T.C., Schriefers, H. & A.D. Friederici, A.D. (2001), Syntactic parsing and working memory: the effects of syntactic complexity, reading span and concurrent load. *Language and Cognitive Processes,* 16 (1): 65-103.

Waters, G. & D. Caplan (1996), The capacity theory of sentence comprehension: critique of Just and Carpenter (1992), *Psychological Review, 4.* 761-772.

Waters, G. & D. Caplan (2005), The relationship between age, processing speed, working memory capacity and language comprehension, *Memory,* 13 (34): 403-413.