

关于日语工作记忆阅读广度 测量方法的研究

马悦 周宜澄 张心雨 常敏 杨泽鹏 韩晓

东南大学，南京

摘要 | 达纳曼和卡朋特（Daneman & Carpenter, 1980）的工作记忆理论为工作记忆广度的测量提供了理论依据，在此基础上出现了众多版本的阅读广度测量方法。但这些测量方法的信效性还值得进一步讨论。本研究通过对 47 名中高水平日语学习者的实证性研究，深入探讨了对材料句子的语义理解考查环节的有无对测试效果的影响这一关键问题。实验结果表明了在阅读广度测试中对材料句子的语义理解进行强制性约束的必要性，同时也证明了笔者们开发的阅读广度测量方法的可信性和有效性。

关键词 | 工作记忆；工作记忆广度；日语阅读广度测量

Copyright © 2022 by author (s) and SciScan Publishing Limited

This article is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/). <https://creativecommons.org/licenses/by-nc/4.0/>



1 引言

1.1 工作记忆容量与阅读理解能力

自从米勒（Miller, 1960）首次提出工作记忆（working memory）一词后，其

基金项目：“国家级大学生创新创业训练计划项目（202110286103）。

作者简介：马悦，周宜澄，张心雨，常敏，杨泽鹏，均为东南大学外国语学院日语系本科生；韩晓，教育学博士，副教授，主要研究方向为日语教育学、认知心理学。

文章引用：马悦，周宜澄，张心雨，等. 关于日语工作记忆阅读广度测量方法的研究[J]. 语言学, 2022, 4(1): 39–53.

<https://doi.org/10.35534/lin.0401005>

被广泛应用于心理学领域研究^[1]。根据巴德利（Baddeley, 2003）对于工作记忆的定义，其是一个容量有限的系统，既有处理功能，又有存储功能。工作记忆的一般作用在于感知新信息、提取长时记忆中存储的知识、并建立思维过程的基础支撑结构^[2]。

在人类大部分认知活动中，工作记忆既负责信息存储任务，又负责信息加工任务。因此，为了解释处理和存储功能这二者的关系，达纳曼和卡朋特在1980年提出共享资源理论。这一理论认为处理和存储依赖于一个共享的资源^[3]。处理需求的增加应该导致能够存储的记忆材料数量的减少，而记忆材料数量的增加应该导致处理效率的下降。简而言之，资源共享观认为，工作记忆中处理和存储的能力相互制衡。

其次，基于工作记忆系统容量有限，对其容量的测量被称为工作记忆广度任务（working memory span），测量方法可以分为听力广度测试（listening span test）、阅读广度测试（reading span test）和运算广度测试（operation span test）。本文着重讨论阅读广度测试方面。阅读理解、推理作为人类高级认知活动，与工作记忆密切关联。1980年，达纳曼和卡朋特基于工作记忆存储以及信息处理资源共享理论，提出了最初的阅读广度测试方法，自此这一测量方法长期以来在心理语言学、二语习得领域的研究中被广泛借鉴。在此基础上，针对不同语种的各种阅读广度测试方法陆续提出。

工作记忆主要由语音回路、视觉空间板、中央执行器和情境缓冲器（Baddeley, 2007）四个部分组成，其中语音回路可以短时间内加工语言和其他听到的声音，包括自己发出的声音^[4]。即便在不出声说话，例如默读语篇时，语音回路仍然处于激活状态。因此，有研究发现在工作记忆测试中，工作记忆广度，尤其是语音回路测验的表现，与阅读能力相关（Bayliss et al., 2005）^[5]。此外，大量研究表明，工作记忆容量可以很好地预测阅读理解能力，哈林顿和索耶（Harrington & Sawyer, 1992）最早发现工作记忆容量高的学习者阅读理解能力更强^[6]。在贾斯特和卡朋特（Just & Carpenter）于1992年提出的句意理解理论中，研究也表明工作记忆能够影响言语阅读理解能力，工作记忆容量能反映阅读理解的成绩^[7]。

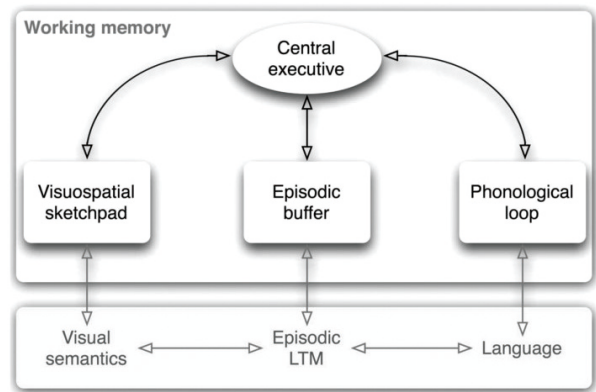


图 1 工作记忆模型 (Baddeley 2000^[8])

Figure 1 The current multi-component model of working memory

1.2 工作记忆阅读广度的测量

许多研究都表明工作记忆容量与阅读理解水平之间存在很大关系。达纳曼和卡朋特也强调工作记忆容量的差别是个体阅读能力差别的主要来源，但传统的短期记忆测量方法，如运算广度、数字广度等与阅读能力关联并不紧密。因此，达纳曼和卡朋特借鉴已经存在的测量方法设计了工作记忆容量的阅读广度测试。工作记忆阅读广度测试可以测量工作记忆存储和加工能力的差异，将其结果与阅读理解测验结果进行相关性分析，便可归纳出工作记忆容量对阅读理解能力方面的影响。以达纳曼和卡朋特设计的工作记忆阅读广度测试为原始版本，国内国外的学者相继开展了关于工作记忆阅读广度的各项实验与研究。在此按照时间顺序对此做一个简单的纵向梳理。

1.2.1 母语阅读广度测量方法

达纳曼和卡朋特 (Daneman & Carpenter, 1980) 借鉴传统短期记忆测量方法，设计了关于工作记忆容量的阅读广度测试。同时进行了相关的阅读理解测试。测试共包括 60 个长度为 13-16 词，句意彼此无关，且句尾单词不重复的句子。阅读广度从 2 句递升至 6 句，每个广度下有 3 组句子，每组句子的数量与该组的阅读广度一致。被试按照自己的语速大声朗读句子，并记住句子的末尾单词。每朗读完一组句子，被试需要按顺序回忆出句末单词。当被试无法正确回忆出

单词或者回忆出的单词顺序错误时,测试停止。三组句子能正确回忆出两组,这三组句子所对应的阅读广度即为被试的阅读广度^[3]。

沃特和卡普兰(Water & Caplan, 1987)对原版达纳曼和卡朋特的阅读广度测试方法着重在测试结果呈现方面进行了改良。测试过程由计算机控制,被试的测试结果由单一末尾单词的回忆数量改为由三项数据共同决定:判断句子语义、句法正误的平均反应时间、在句子判断中产生错误的数量以及句末单词被错误回忆的次数。三项数据的复合得分为阅读广度测试结果^[9]。

国内关于汉语阅读广度测量的研究。鲁忠义和范宁(2006)使用计算机进行实验。每个句子完整出现6秒,被试需完成句意理解及句末词语记忆。6秒过后原句消失,被试需判断与原句相关的一个短句是否符合原句句意。计分方式与测试结果呈现仍采用达纳曼和卡朋特的原始方法^[10]。

杨奇伟(2007)以原始版本为基础,并对其进行改良。使用计算机呈现试题,被试同样需要朗读句子并记住句末词语(双音节汉字词语)。一组句子展示完毕后被试需按顺序回忆出词语,回忆不出或错误则测试终止。但值得注意的是,杨奇伟对句子的呈现方式略作调整,为了防止被试直接记忆句末词语而忽略句意理解,将会先呈现缺少句末词汇的句子,句末词汇则在一秒后出现^[11]。

1.2.2 二语阅读广度测量方法

二语习得,即第二语言习得,一般被认为是除母语习得之外的任意语言的学习。自从麦克劳克林和斯凯恩(McLaughlin & Skehan)等学者将工作记忆的概念引入二语习得领域之后,引起了该领域部分学者的浓厚兴趣。阅读广度测量范式可以评估个人工作记忆的信息存储与加工能力差异,将之与二语习得领域相关研究结果进行数据分析,便可归纳出工作记忆容量与二语习得某一领域的关系(温植胜,2007)^[12]。

哈林顿和索耶仿照达纳曼和卡朋特的阅读广度测试设计了针对以日语为母语的英语学习者的阅读广度测试。测试包括42个在长度和语法上都更为简单的句子,这些句子被分成若干组,被试需要大声朗读句子,并在每组句子朗读完毕后回忆并写下句末单词。同时该测试也包括语法判断测试来尽可能确保被试按要求进行句子理解。测试分数仅包括能正确回忆出的单词总数^[6]。

奥萨卡 (Osaka, 1992) 设计了阅读广度测试以进一步研究工作记忆容量与母语 (L1) 和第二语言 (L2) 之间的关系, 从而判断工作记忆容量是否与语种有关。测试对象为以日语为母语的英语学习者。被试在进行达纳曼和卡朋特的阅读广度测试的同时进行日语版本的阅读广度测试和以英语为第二语言 (ESL) 的阅读广度测试。被试被要求大声朗读每组句子, 并在每组句子朗读完毕后口头回忆出句末单词。此外不设置附加的句子判断测试。测试结果为被试能连续正确回忆出的至少三组句子末尾的单词数量^[13]。

国内研究目前主要针对汉语为母语的英语学习者开展。张晓东 (2014) 以国内某高校英语专业学生为受试, 探讨短时记忆、工作记忆与词汇知识对二语阅读能力的影响, 认为工作记忆容量对二语阅读水平具有显著预测力^[14]。

倪锦诚 (2017) 则以非英语专业的本科生和硕士生为被试, 进行了阅读广度测试和阅读能力测试, 在阅读广度测试中, 要求被试判断与原句相关的短句是否符合原句句意, 从而尽可能保证被试按要求进行句意理解。通过相关性分析, 得出工作记忆容量与英语阅读准确度显著相关的结论, 但其与英语阅读加工时效不存在显著相关性^[15]。

1.2.3 工作记忆阅读广度测量方法的探讨

通过对以上版本阅读广度测试的分析比较, 可以看到继达纳曼和卡朋特设计了最早的阅读广度测试版本以来, 后续的阅读广度测试都以此版本为基础, 不断进行修改和完善, 其重点在于如何尽可能地确保被试按要求理解句子, 即保证工作记忆的加工环节。部分研究者认为被试即便在测试中进行了句子的朗读, 但这不能保证其一定实施了语义加工和语义理解, 因此主张在原始版本的测试基础上增加句意判断的环节。但增加句意判断环节的改良版本的测试和不进行句意判断的测试之间, 其测试效果孰优孰劣, 尚缺少明确的研究结论。

另一方面, 即使是主张增加句意判断环节的研究者中, 对于选择何种具体的句意判断测试方法也有争议。目前现有研究中句意判断的方式主要有两种: 一是判断与原句相关的独立短句是否与原句句意相符, 二是判断原句语法句法是否正确。另外还需要指出的是, 增加句意判断环节, 确实可以在一定程度上确保被试按要求进行句意理解, 但两种方式同时也都存在一定不足。因为独立

短句与原句相关,因此判断独立短句是否符合原句的方式容易导致原句内容的二次记忆。能否正确判断语法句法则受到被试本身语言水平的影响,在二语阅读广度测试中容易混淆二语水平因素与工作记忆因素对阅读理解的影响^[16]。一些固定的语法句法规则即使不依靠句意理解也同样能够判断。因此,在具备句意判断环节的二语阅读广度测试的开发中,如何克服上述缺陷是研究的重点和难点。松见法男、福田伦子、古本裕美等(2009)面向日语学习者开发的听力广度测试(listening span test)中采用了一种独特的方法,他们使用的句子在语法句法规则上没有问题,但有一半句子的句意存在逻辑上的问题^[17]。比如,“財布はお菓子を入れるものである(钱包是用来装点心的)”这个句子表达的意思跟普通人的生活认知是相违背的,从正常逻辑上是不成立的。他们通过让被试在测试中针对听到的句子一边记忆目标单词,一边判断该句子的逻辑是否成立,从而保证被试在测试中进行句意理解。由于这种方法没有增加额外的句意判断的短句,也就避免了笔者在上文中指出的相应问题。但这种方法是否同样适用于阅读广度的测量,尚未可知,还需进一步验证。

以上研究表明,工作记忆阅读广度确实与阅读理解能力存在显著的相关性。但同时可以发现,尽管工作记忆容量的相关研究在二语习得领域蓬勃发展,但国内研究目前仍以二语为英语的研究为主,针对汉语母语者的日语学习的相关研究较少。可以说,目前阅读广度测试特别是二语阅读广度测试仍处于不断探索的阶段。

1.3 研究目的

基于目前国内工作记忆阅读广度测量工具的研究现状,针对以汉语为母语的日语学习者,以是否包含句意判断环节为区分,我们开发了两版略有差异的日语版本工作记忆阅读广度测量工具。本实验旨在通过工作记忆阅读广度测试以及日语阅读理解测验,证明工作记忆阅读广度与日语阅读理解水平之间存在显著相关性,并以此验证该版本日语阅读广度测量工具的有效性。其后对两个测试版本所得数据分别进行相关性分析,对比出更为有效的日语阅读广度测量工具,探究以判断句意是否符合逻辑的方式进行句意判断是否有助于更为准确

的测量阅读广度。

近年来,在国内核心期刊上发表的工作记忆容量相关研究论文在二语习得领域大多针对不同年龄层汉语母语者的英语学习,较少涉及其他语言,研究内容也大多从同声传译、语法学习、听力等角度出发,关注阅读理解能力的研究相对较少^[18]。本研究则面向国内广大日语学习者,探究日语工作记忆阅读广度与阅读理解能力之间的关系,在一定程度上是对国内二语习得领域工作记忆容量相关研究的有益补充。面对国内日语学习者数量不断增加的近况,我们也希望能将本研究开发的日语版工作记忆测量工具应用于日语教学实践,促进国内日本语教学更加科学高效地开展。

2 研究方法

2.1 研究对象

研究对象为南京某高校日语专业大二至大四年级的本科生共47人,其中女生29人,男生18人。最终共47人参与并完成了日语阅读广度测试以及阅读理解测试,为有效被试,有效率为100%。所有研究对象均为以汉语为母语的日语学习者,均没有赴日交换留学经历,平均日语学习年限为2.5年,视为中高水平日语学习者。且同一年级本科生日语水平相近,具有相似的日语学习环境和学习经历。

2.2 实验材料

2.2.1 阅读理解测验

根据被试日语水平的差异设计了两套测验试题,分别为A卷和B卷。两套试卷题目均选自历年日本语能力考试(JLPT)以及使用日本语鉴定考试(J.TEST)真题,题型均为阅读选择题。A卷的难度水平相当于日本语能力考试N3至N2等级,为大二年级本科生阅读理解测验试题。B卷的难度水平相当于日本语能力考试N1等级,为大三、大四年级本科生测验试题。

A卷共包括4篇300字左右的日语短篇阅读,并设有10道选择题形式的阅

读理解题。B 卷共包括 2 篇 300–400 字左右的短篇阅读、1 篇 600 字左右的中篇阅读以及 1 篇 1000 字左右的长篇阅读，同样设有 10 道选择题形式的阅读理解题。

2.2.2 日语 RST 测试

基于本研究的研究目的，笔者设计了两套略有不同的日语阅读广度测量工具，分别为 RST1 和 RST2。其中 RST1 的测试方法参考了达纳曼和卡朋特的原始版本，RST2 则参考松见法男、福田伦子、古本裕美等（2009）的听力广度测试，RST1 和 RST2 的主要区别在于是否包含句意理解测试环节。两个版本的 RST 测量工具均包括 46 个长度为 8–20 词的句子，且两个版本所含句子均不重复。其中前四个句子为练习题，不计入测试结果，后 42 个句子为正式测试题。句子中出现的所有单词均为日语 N2 及以下水平，且句中实词不重复，句意内容均为一般性常识内容。经过测试后对被试进行确认，RST1 和 RST2 中的所有日语单词和句式均已学习，不存在未知内容。两个版本的测量工具的阅读广度均从 2 句递升至 5 句，每个广度下包含 3 组句子，每组句子数量与其阅读广度水平相对应（如：2 句水平下有 3 个句组，每组中又分别包含 2 个句子。2 句以上水平以此类推。）

RST1 中要求被试记忆的目标单词（加粗字体，使其与其他单词区分）均为长度 3–6 拍的名词，且目标单词在句中位置不固定。RST1 中使用的测试句示例如表 1 所示。

表 1 RST1 中所用材料句及分组示例（2 句水平）

Table 1 Material sentences and group examples in RST 1 (Level 2)

		目标单词
句组 1	句 1: 一ヶ月前にホテルの 予約 をした。	予約
	句 2: ラジオ の音が大きいので小さくした。	ラジオ
句组 2	句 1: 両親と相談してから 結婚を決める 。	決める
	句 2: 仕事で疲れたので いつも より早く寝た。	いつも
句组 3	句 1: 簡単な 漢字 が書けるようになった。	漢字
	句 2: 彼女にあげる プレゼント を買いに行った。	プレゼント

RST2 中使用的 42 个实验测试句中，符合逻辑的有 19 句，逻辑不符的

有 23 句。要求被试记忆的目标单词（加粗字体，使其与其他单词区分）均为长度 3–6 拍的名词，且目标单词固定于句首。RST2 中使用的测试句示例如表 2 所示。

表 2 RST2 中所用材料句及分组示例（2 句水平）

Table 2 Material sentences and group examples in RST 1 (Level 2)

		判断结果	句首单词
句组 1	句 1: 子ども は大人よりも若い。	正	子ども
	句 2: 財布 はお菓子を入れるものである。	误	財布
句组 2	句 1: 大学 には研究室がある。	正	大学
	句 2: サンドイッチ を作るためにはパンが必要だ。	正	サンドイッチ
句组 3	句 1: 一週間 は 31 日である。	误	一週間
	句 2: 魚 は湖にしかない。	误	魚

2.2.3 研究程序

实验以小组为单位开展，每组 2–3 人，均为相同年级。被试按照先阅读理解测试后 RST 测试的顺序参加了测试。

（1）阅读理解能力测试及评分

测验形式为纸质试卷，针对大二年级的 A 卷测验时间为 30 分钟，针对大三、大四年级的 B 卷测验时间为 25 分钟，均允许提前交卷，到达规定时间后由实验者回收试卷。等到一组被试全部完成第一项阅读理解测验之后，开始 RST 测试。

阅读理解能力测试 A 卷和 B 卷均包含 10 道选择题，每题 10 分，满分 100 分。

（2）RST 测试及评分

包括 RST1 和 RST2 两部分，同样采用纸质试卷形式。

在 RST1 中，被试从第一页开始按顺序朗读句子，理解句意的同时记忆目标单词，每个句子仅读一遍，且在朗读过程中不允许向前或向后翻页。该组全部句子朗读完毕后迅速翻页，在下一页给出的括号中按顺序填入记忆的目标词语。完成后翻至下一页开始下一组句子的朗读和记忆，以此类推，直到完成全部测试内容。

在 RST2 中，被试同样按顺序从第一页开始朗读句子，同时进行句意理解并

记忆目标单词,朗读过程中不允许随意翻页。该组全部句子朗读完毕后迅速翻页,在下一页给出的横线上写出对句意是否符合逻辑的判断结果,符合逻辑用圆圈表示,逻辑不符用打叉表示。完成判断后翻页,在下一页给出的括号中按顺序填入记忆的目标词语。完成后翻至下一页开始下一组句子的朗读和记忆,以此类推,直到完成全部测试内容。

为了防止顺序效应影响实验结果,顺序相邻的被试之间采用了平衡技术(counter balance),即前一位被试按 RST1 到 RST2 顺序测试,则后一位被试按 RST2 到 RST1 的顺序测试。被试完成全部两项 RST 测试后即完成全部实验。

RST1 相同水平的 3 个测试组中,被试每正确通过其中的 2 组或以上则视为完成了该水平的测试,可获得该水平代表的分值——2 句水平得 2 分,3 句水平得 3 分,以此类推——并继续进入后一水平的测试。相同水平的 3 个测试组中,被试若只能正确通过其中的 1 组,则测试结束,被试可额外获得 0.5 分的附加分,即被试的 RST 最终得分为前一测试水平的分值与附加分(0.5 分)之和(如:3 句水平测试中完成了 1 组,得分为 2.5 分)。相同水平的 3 个测试组中 1 组也不能通过的,RST 得分为已通过的前一测试水平的分值,无附加分。RST2 的评分方式与 RST1 基本相同,但与 RST1 中被试只需要正确再现目标单词即可视为通过的评价标准不同,在 RST2 中只有当被试对组内各个句子的单词记忆和逻辑判断都正确时才被视为通过,有一项错误则视为该组不通过,不能得分。RST1 和 RST2 的分值区间均为 2-5 分。其得分越高,则表示该被试的工作记忆容量大,反之则代表被试工作记忆容量小。

3 研究结果及讨论

47 名被试的阅读理解能力测试的平均得分为 76.65 分,标准差为 16.82, RST1 的平均得分为 2.77 分,标准差为 0.67, RST2 的平均得分为 2.34 分,标准差为 0.84。使用 SPSS 对被试的阅读理解能力测试成绩和 RST1 成绩、RST2 成绩分别进行相关分析(本研究的显著性水平均设定为 5%),结果显示:被试的阅读理解能力测试成绩和 RST1 测试成绩微弱相关($r=0.288$,

$p=0.132$), 另一方面, 被试的阅读理解能力测试成绩和 RST2 测试成绩之间的相关性显著 ($r=0.636$, $p<0.01$)。

相关分析的结果说明基于达纳曼和卡朋特的经典的阅读广度测量方法开发的 RST1 的测试成绩与中高级水平日语学习者的日语阅读理解能力之间的相关性较弱, RST1 的测试成绩不能很好地预测学习者的日语阅读理解能力, 与张晓东 (2014) 等研究的结论相左。与此相对, 基于松见法男、福田伦子、古本裕美等 (2009) 的听力广度测试开发的 RST2 的测试成绩中高级水平日语学习者的日语阅读理解能力之间存在较强的相关性, 说明 RST2 的测试成绩对学习者的日语阅读理解能力有显著的预测力。也就是说, 相比 RST1, RST2 是更理想的阅读广度测量方法。

另一方面, 上述结果还说明在二语的阅读广度测试中, 如果不对语义理解进行某种形式的强制性约束, 则被试很可能为了取得更好的测试成绩, 会把更多注意力放到单词的记忆上, 而不是语义加工的过程中, 而这会使测试结果不能准确反映被试的阅读理解能力。迫田久美子 (2010) 指出外语学习者的注意力分配策略受其外语水平的高低影响^[19]。因此笔者推断在 RST1 测试中, 日语熟练度高的高水平学习者和日语熟练度相对较低的中水平学习者的注意力分配方式很可能有所不同。为了验证这一点, 笔者尝试将被试按照学年分为二年级组、三年级组和四年级组, 然后针对各个组重新进行了相关分析。结果如表 3 所示。

表 3 按年级分组后的相关分析结果

Table 3 Relevant analysis results after dividing groups by grades

	RST1	RST2
二年级组	$r=0.413^*$	$r=0.641^*$
三年级组	$r=0.201$	$r=0.712^{**}$
四年级组	$r=0.197$	$r=0.663^{**}$

注: $*p<0.05$, $**p<0.01$ 。

以上的分析方法也许并不十分严谨, 但也可以从一定程度上反应出不同日

语水平的被试,尤其是中水平的二年级组和高水平的三、四年级组相比,在 RST1 的测试结果上存在较为明显的差异。导致这种差异产生的原因很可能就在于 RST1 中没有对语义理解进行强制性约束。换言之,对二语水平较高的被试而言,其朗读等语音加工的过程中未必同步伴有语义理解,即语义加工过程。与之相对的中水平或更低水平的被试,则受其二语熟练程度的制约,他们的语义加工过程很难脱离语音加工过程独立存在,因此与高水平被试相比,他们的 RST1 的成绩反而对自身的二语阅读理解能力有着更显著的预测力。这一结论也进一步印证了前文的结论,即增加了语义理解考查环节的 RST2 比缺少了语义理解考查环节的 RST1 的测试效果要更为理想,也更能体现被试的二语阅读理解能力。

4 结语

本研究基于达纳曼和卡朋特(1980)的工作记忆理论,对目前存在的工作记忆阅读广度测量方法进行了改良和完善,设计开发出了面向日语学习者的日语版阅读广度测量方法。并且,以47名中高水平日语学习者作为被试,通过日语阅读广度测试成绩与学习者阅读理解能力成绩之间的相关分析,探讨了阅读广度测试过程中对材料句子的语义理解考查环节的有无对阅读广度测试的有效性的影响情况。实验结果显示,笔者们开发的具备对材料句子的语义理解考查环节的日语阅读广度测试(RST2)成绩和被试的阅读理解能力测试成绩之间的相关性显著,而不考查被试对材料句子的语义理解的日语阅读广度测试(RST1)成绩和被试的阅读理解能力测试成绩之间的相关性不显著。该结果既说明了在阅读广度测试中对材料句子的语义理解进行强制性约束的必要性,也证明了笔者们开发的阅读广度测量方法(RST2)的可信性和有效性。

由于条件限制,本研究还存在以下两方面的问题:一是,不同日语水平的被试人数的配比方面还有一些问题(比如,三四年级高水平被试人数比重较大),实验结果尚不十分完备。二是,本研究的 RST2 主要依托松见法男、福田伦子、古本裕美等(2009)的听力广度测试,并没有对其测试材料,尤其是材料中的

目标单词的长短、学习者的熟悉程度等方面做过多的讨论。关于以上两方面的问题,将在今后的研究中做进一步的深入探讨。

参考文献

- [1] Brown R. Plans and the Structure of Behavior by George A. Miller; Eugene Galanter; Karl H. Pribram [J]. *Language*, 1960, 36 (4): 527-532.
- [2] Baddeley A. Working memory and language: An overview [J]. *Journal of Communication Disorders*, 2003, 36 (3): 189-208.
- [3] Daneman M, Carpenter P A. Individual differences in working memory and reading [J]. *Journal of Verbal Learning & Verbal Behavior*, 1980, 19 (4): 450-466.
- [4] Baddeley A D. Working Memory, Thought, and Action [M]. Oxford: Oxford University Press, 2007: 2-15.
- [5] Bayliss D M, Jarrold C, Baddeley A D, et al. Differential constraints on the working memory and reading abilities of individuals with learning difficulties and typically developing children [J]. *Journal of Experimental Child Psychology*, 2005, 92 (1): 76-99.
- [6] Harrington M, Sawyer M. L2 Working Memory Capacity and L2 Reading Skill [J]. *Studies in Second Language Acquisition*, 1992, 14 (1): 25-38.
- [7] Waters G S, Caplan D. The capacity theory of sentence comprehension: Critique of Just and Carpenter (1992) [J]. *Psychological Review*, 1996, 103 (4): 761-772.
- [8] Baddeley A. The episodic buffer: a new component of working memory? [J]. *Trends in cognitive sciences*. 2000, 4 (11): 417-423.
- [9] Waters G, Caplan D, Hildebrandt N. Working memory and written sentence comprehension [J]. *The Psychology of Reading*, 1987, 9: 531-555.
- [10] 鲁忠义, 范宁. 工作记忆广度与汉语句子语境效应的关系 [J]. *心理学报*,

- 2006, 38 (1) : 22-29.
- [11] 杨奇伟. 英、汉语阅读广度的影响因素: 工作记忆广度理论的潜变量分析 [D]. 西北大学, 2007.
- [12] 温植胜. 外语学能研究新视角——工作记忆效应 [J]. 现代外语, 2007 (1) : 87-95+110.
- [13] Osaka M, Osaka N. Language-independent working memory as measured by Japanese and English reading span tests [J]. Bulletin of the Psychonomic Society, 1992, 30: 287-289.
- [14] 张晓东. 短时记忆、工作记忆及词汇知识对二语接收性言语技能的影响 [J]. 外语界, 2014 (5) : 38-47.
- [15] 倪锦诚. 工作记忆容量对二语阅读能力的影响研究 [J]. 解放军外国语学院学报, 2017, 40 (3) : 79-85+160.
- [16] Juffs A, Harrington M. Aspects of working memory in L2 learning [J]. Language Teaching, 2011, 44 (2) : 137-166.
- [17] 松見法男, 福田倫子, 古本裕美, 邱俞瑗. 日本語学習者用リスニングスパンテストの開発—台湾人日本語学習者を対象とした信頼性と妥当性の検討— [J]. 日本語教育, 2009, 141: 68-78.
- [18] 陈士法, 崔涛涛, 罗小姝. 国内外语界工作记忆与第二语言习得关系研究 20 年 [J]. 北京第二外国语学院学报, 2016, 38 (2) : 15-27+132.
- [19] 迫田久美子. 日本語学習者に対するシャドーイング実践研究—第二言語習得研究に基づく運用力の養成を目指して— [J]. 第二言語としての日本語の習得研究, 2010, 13: 5-21.

A Study on the Japanese Working Memory Reading Span Test

Ma Yue Zhou Yicheng Zhang Xinyu Chang Min
Yang Zepeng Han Xiao

Southeast University, Nanjing

Abstract: Working memory theory proposed by Daneman and Carpenter in 1980 provided the theoretical basis for the measurement of working memory span. Based on it, various versions of reading span tests appeared. However, reliability and validity of these methods worth further discussion. Through an empirical study of 47 intermediate and high level Japanese learners, this research gives an insight into the key issue that whether semantic comprehension sector on material sentences have influence on test effects. The results demonstrate the necessity of mandatory constraints on the semantic comprehension of material sentences in reading span test. Meanwhile, the experiment also proves that the credibility and validity of the reading span test developed by authors.

Key words: Working memory; Working memory span; Japanese reading span test