

## 科学研究費助成事業 研究成果報告書

平成 26 年 6 月 10 日現在

機関番号：11301

研究種目：若手研究(B)

研究期間：2012～2013

課題番号：24790606

研究課題名(和文) 語彙力を測定するアセスメントツールの開発ー出生コホートでの活用を目指してー

研究課題名(英文) Development and Validation of Japanese version of the Boston Naming Test

研究代表者

龍田 希 (TATSUTA, Nozomi)

東北大学・医学(系)研究科(研究院)・助教

研究者番号：40547709

交付決定額(研究期間全体)：(直接経費) 3,400,000円、(間接経費) 1,020,000円

研究成果の概要(和文)：疫学調査で子どもの言語発達を把握するために、米国で開発されたBoston Naming Testが用いられることが多いが、ここには日本人には馴染みのない語彙が含まれている。日本人の語彙力を調べるために、日本の社会文化的背景を考慮した検査を開発することとした。教育基本語彙の基本的研究(明治書院, 2009)に記載のある語彙から300語を抜粋し、そのイメージする絵を白黒の線画で描画した。それを5歳～12歳の391名に提示し、その名称を回答させた。回答結果を用いて、信頼性(内的整合性、折半法)と妥当性(併存的妥当性、内容的妥当性)を調べ、条件を満たした60項目を日本語版語彙検査の項目とした。

研究成果の概要(英文)：The Boston Naming Test (BNT) has been used to assess the language development of children in many epidemiology studies, and its usefulness is confirmed. The contents of the items should alter because many items in the BNT were considered inapplicable to the Japanese. The purpose of this study was to develop and evaluate the reliability and validity of the Japanese version of BNT. The participants of the study consisted of 391 children (aged 5-12 years). The Kuder-Richardson Formula 20 of internal consistency was 0.936 and spearman-Brown split-half reliability coefficients was found as 0.943. The significant correlation between the score of Japanese version of BNT and original BNT for concurrent validity was examined ( $r=0.918$ ,  $p=0.001$ ). For content validity, using point biserial correlation, examined relationships between each item score and total score. The Japanese version of the BNT has good reliability and validity, and it can measure the language ability of Japanese children.

研究分野：医歯薬学

科研費の分科・細目：社会医学・公衆衛生学・健康科学

キーワード：語彙検査 子ども 疫学調査

## 1. 研究開始当初の背景

疫学調査で子どもの言語発達を把握するために、米国 (Boston) で開発された Boston Naming Test (BNT) を用いることが多い。これは、提示された絵の名称を回答するという語彙力(ごい力 = ことばの名称をどの程度知っているか)を調べる検査である。BNTは、調査にかかる時間も15分程度と短く、また結果の評価方法も一般の発達検査や知能検査と比較しても複雑ではない。そのため、対象児や検査者への負担も少なく、また多くの子どもを対象とする疫学調査では有用である。ところが、BNTには日本人には馴染みのない語彙(例えば、「イグルー」や「プレッツェル」など)が含まれていたり、日本で使用されるものとは異なる形をした語彙(例えば、「巻物」は上下に心棒があり縦に開くものが描かれているが、これは英文が左から右に書き進めることと関連している。和文は上から下に書き進めるため、左右に心棒があり横に開くものを想像する)が含まれているため、日本人がこれを実施すると、米国人を対象として標準化されたBNTの正答率とは差異がみられる。諸外国では、BNTを基盤としてそれぞれの国に適した語彙検査を開発している。我が国には、標準化された語彙検査として、PVT-R 絵画語い発達検査(日本文化科学社, 2008)があるが、これは複数の絵の中から検査者が呼称した絵を指差すという検査であり、BNTを基盤とした語彙検査は存在しない。

そこで、疫学調査での使用を想定した日本の社会文化的背景に適した日本語版語彙検査を独自に開発する必要がある。

## 2. 研究の目的

海外で進められている多くの疫学調査で子どもの語彙力を把握するために、BNTが用いられており、その有用性が報告されている。このBNTの得点は、社会文化的背景の影響を大きく受けることがわかっていることから、日本人を対象とする場合には、日本の社会環境に適した語彙検査が必要である。そこで本研究では、疫学調査で子どもの語彙力を評価するための語彙検査を開発し、その信頼性および妥当性を検証することとした。

## 3. 研究の方法

### (1) 日本語版語彙検査の項目選定

日本語版語彙検査の項目を選定するために、日本で標準化されているPVT-R 絵画語い発達検査のマニュアルを参考にした。PVT-R 絵画語い発達検査では、教育基本語彙の基本的研究(明治書院, 2001)に掲載された語彙からランダムに選択された。本研究でもPVT-Rを参考に教育基本語彙の基本的研究・増補改訂版(明治書院, 2009)から日本語版語彙検査の項目を抜粋することとした。

BNTは60項目から構成されており、対象児への負担等も考慮すると60項目適度が適

当であると思われた。そこで日本語版語彙検査でも最終的に60項目を選択する計画とした。教育基本語彙の基本的研究・増補改訂版にある語彙から抜粋された語をイメージする絵を描画するが、検査の信頼性および妥当性を検証する中で、項目が減少することが予想されたことから、300語を抽出した。

### (2) 絵の作成

BNTで使用されている絵は、白黒の線で描かれていることから、本研究においても白黒の線で描くこととした。絵は、筆圧や絵の雰囲気などの相違を考慮し、一人の担当者がすべてを描くこととした。

### (3) 成人調査

描画した300枚の絵について、我々が意図する回答が得られるかを調べるために、一般成人を対象に調査を実施した。20代~60代の一般成人22名(男性11名、女性11名)に絵を提示し、それぞれの名称を回答させた。ここで得られた結果を基に、絵を修正したり、書き直したりした。その場合にはそれらの項目のみ再度調査を実施した。

### (4) 子ども調査

成人調査で選択された絵を実際に5歳から12歳の子どもに提示し、その名称を回答させる調査を実施した。調査は391名を対象に実施したが、母国語を日本語としない子どもや障害のある子どもが含まれたため、日本語を母国語とする健常児に限定し、386名(男児213名、女児173)を解析の対象とした。

調査の内容は、①成人調査で選択された日本語版語彙検査の項目候補と②BNTの2課題であった。386名のうち94名については知能検査(Wechsler Intelligence Scale for Children Third edition, WISC-III)もあわせて実施した。日本語版語彙検査については、iPadを用いて絵を提示した。

調査に同伴した対象児の母親には、養育者の学歴や喫煙状況、飲酒の有無、対象児の睡眠時間等の家庭での生活環境に関する質問票への回答を依頼した。

### (5) 信頼性と妥当性の検証

子ども調査で得られた結果を用いて、日本語版語彙検査の信頼性と妥当性を調べた。信頼性とはどれだけ正確で安定した検査であるかを調べる方法であり、妥当性とは目的にあったことを測定できているかを検証する方法である。

信頼性の検証には、各項目が同じもの(ここでは語彙力)を測定しているかを調べるために内的整合性を調べた。その手法には、Kuder-Richardsonの公式20(以下、KR-20)を用いた。これは、より $\pm 1$ に近い数字であれば信頼性が確保されたとみなされる。次に、検査項目の安定性を調べるために1つの検査に含まれる項目を2つに分割して(本研究で

は奇数番号と偶数番号で分割した)、得点の一致率を調べた(折半法)。その手法には、Spearman-Brownの公式を用いた。これは、相関係数と同じように $\pm 1$ に近いほど一致率が高いとみなされる。なお、本研究では再検査信頼性は実施しなかった。その理由は、再検査信頼性はいつ調査を実施しても同じ結果が得られることを調べる方法であるが、語彙検査では学習効果により、二回目のほうで得点が高くなることが予想されたためである。

妥当性の検証には、基準関連妥当性の中の併存的妥当性を調べた。併存的妥当性とは、外部の基準に照らし合わせたとき、どの程度関連性があるかを調べることを指す。本研究では、日本語版語彙検査と同時にBNTも実施しており、それらの得点間の関連性をPearsonの積率相関係数から調べた。また、一部の対象児にはWISC-IIIも実施しており、これらとの関連性もPearsonの積率相関係数から調べた。WISC-IIIとの比較については、それぞれの下位検査の粗点を解析に用いた。WISC-IIIで算出される知能指数(Intelligence Quotient, IQ)は同年齢集団内での位置を算出するため、ある年齢集団において平均的な得点をとった場合にIQが100となるように換算される。日本語版語彙検査では年齢によって平均得点が変わるので、WISC-IIIとの比較では下位検査の粗点を用いた。次に、基準関連妥当性の内容的妥当性を調べた。内容的妥当性とは、項目が全体から偏りなく測りたいことが測れているかを調べることを指す。その手法として、総得点と各項目の得点との関連性を調べる項目弁別力(点双列相関係数)を算出した。点双列相関係数はより $\pm 1$ に近い場合に妥当性が高いとされる。

以上の統計処理には、SPSS version 19.0を用いた(IBM Inc, Japan)。統計学的有意水準は5%とした。

#### (6) 調査における倫理的配慮について

一般成人を対象とした場合には、対象者本人に研究内容、研究の趣旨を説明し、書面による同意を得た。子どもを対象とした場合には、未成年であることから、保護者に口頭で説明をした上で書面にて同意を得るとともに、対象児に調査の内容を口頭で説明し、同意が得られた場合に調査を実施した。

本研究の実施にあたり、東北大学医学系研究科倫理委員会に研究計画を提出し、承認を得た上で調査を進めた(受付番号2012-1-72)。

## 4. 研究成果

### (1) 成人調査

成人調査では、300枚の絵が日本語版語彙検査の項目として適切かを判断することを目的とした。その中で、我々が意図した回答が得られなかった項目(例えば、「フライパン」を虫眼鏡と回答したり、「シュークリーム」をシ

ルクハットと回答した)等の91項目については項目として妥当ではないと判断し、項目候補から除外した。また、14項目については絵を修正し、再度調査を実施した(例えば、「セイロ」の絵に肉まんの絵が描かれており、肉まんや焼売という回答が見られたため、肉まんの絵を削除した。「指輪」の絵に大きなダイヤモンドが描かれており、ダイヤモンドという回答が見られたため、ダイヤモンドを小さくした)。その上で、209項目を語彙検査の項目候補とした。

### (2) 子ども調査

子ども調査では、成人調査で選定された209項目を用いた。所要時間は15~20分程度であった。回答についてみると、我々が意図した回答が得られなかった項目(例:正答は「マンホール」だが、レモンやピザ、ルンバ等の回答が得られた)や、問題がわかりにくかった項目(例:正答は「浴衣」だが、団扇や女の人、夏祭り等の回答が得られた)、複数の正答が考えられた項目(例:正答は「入れ歯」だが、歯や歯の模型、虫歯等の回答が得られた)が複数あり、絵の質・課題の質が良くないと判断し、項目から除外した。

次に、BNTでは難易度が低い項目から順番に並べられている。日本語版語彙検査においても、項目毎に正答率を算出し、正答率が高い項目(易しい項目)から低い項目(難しい項目)までが網羅されるよう配慮して項目を選定した。

### (3) 信頼性と妥当性の検証

信頼性については、内的整合性を調べた。KR-20を算出した結果、0.936となり1に近い値が得られた。次に、検査項目の安定性を調べた(折半法)。Spearman-Brownの公式によると、0.943となり、こちらも1に近い値が示された。以上の結果より、日本語版語彙検査の信頼性が担保されたと考えられた。

妥当性については、まず併存的妥当性を調べた。日本語版語彙検査とBNTとの関連性を調べたところ、 $r=0.918$  ( $p<0.001$ )であった。日本語版語彙検査とWISC-IIIの言語性IQを測定する下位検査の結果との間にも統計学的に有意な関連が認められた(知識  $r=0.802$ ,  $p<0.001$ ; 類似  $r=0.742$ ,  $p<0.001$ ; 単語  $r=0.731$ ,  $p<0.001$ ; 理解  $r=0.765$ ,  $p<0.001$ )。相関係数に着目すると、いずれも0.7を越えており高い関連性が示された。次に、内的妥当性を調べた。その結果、3項目を除きいずれも統計学的に有意な相関関係が示された。有意ではなかった3項目のうち、1項目は正答率が100%の項目であり、2項目は正答率が著しく低かった項目であった(正答率は1.0%と1.8%であった)。以上の結果より、日本語版語彙検査の妥当性が確認された。

### (4) 日本語版語彙検査結果と関連のある要因

日本語版語彙検査で得られた得点と関連のある要因を対象児の母親を対象に実施した質問票の結果から検討した。まず、対象児の基本属性との関連性を調べた。日本語版語彙検査の得点について、性差は認められなかった ( $t(384)=-0.166, p=0.868$ ) が、年齢が上がるほど得点が高くなった ( $r(386)=0.777, p<0.001$ )。出生順位についてみると、第一子のほうが第一子以降の子どもよりも得点が高く ( $t(384)=5.518, p<0.001$ ) (図1)、弟や妹のいる子どもはいない子どもと比べて得点が高かった ( $t(384)=-5.199, p<0.001$ )。次に、対象児の家庭での生活環境についても調べた。「お子さまは本を読むことが好きですか」という質問に対し、本を読むことが好きな子どもはそうではない子どもと比べて得点が高かった ( $F(2, 385)=6.233, p=0.002$ )。テレビを観る時間は年齢があがるほど短くなるため、年齢で補正すると、テレビを観る時間が長い子どもほど語彙検査の得点が低くなるのがわかった ( $r=-0.169, p=0.001$ ) (図2)。学習塾や勉強に関する通信教育を使用している子どもはしていない子どもと比べて得点が高くなること示された ( $t(384)=-4.877, p<0.001$ ) (図3)。これらの結果から、テレビを観る時間を読書の時間に変えることで語彙力の増加が期待された。

なお、両親の喫煙状況、飲酒状況、家庭の年収、両親の学歴、対象児の睡眠時間や利き手等と語彙検査の得点との間に関連性は認められなかった。

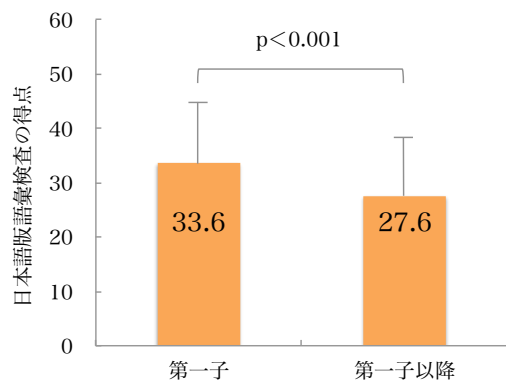


図1 日本語版語彙検査の得点と出生順位

#### (5) 今後の課題

本研究では、日本語版語彙検査を作成し、その信頼性と妥当性を検証することを目的とした。日本人の社会文化的背景に適した日本語版語彙検査を作成し、内的整合性等の信頼性、併存的妥当性や内容的妥当性等の妥当性を検証しており、概ね目標は達成できたと考える。本研究では、東北地方で調査を実施した。今後は他地域の子どもの対象に調査を実施し、本研究の対象集団で確かめられた信頼性や妥当性が他の集団でも同様に成立するかを明らかにする交差妥当性を検証し、日本語版語彙検査の精度を高める計画である。

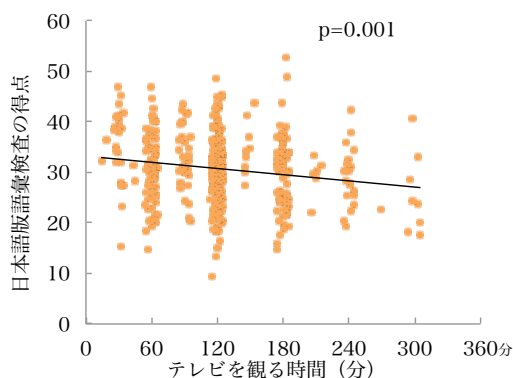


図2 日本語版語彙検査の得点とテレビを観る時間

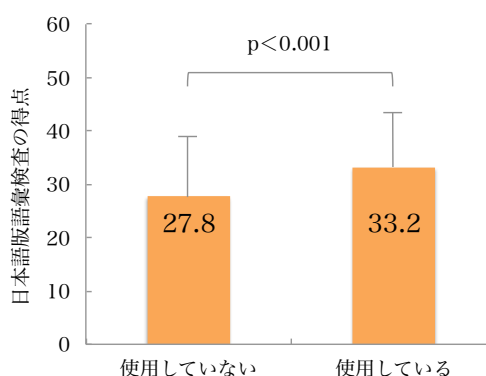


図3 日本語版語彙検査と学習塾使用の有無

## 5. 主な発表論文等

[雑誌論文] (計 1 件)

- (1) 龍田 希、仲井邦彦、鈴木恵太、黒川修行、細川 徹、佐藤 洋、日本人小学 1 年生を対象に実施した Boston Naming Test の標準データの検討、日本衛生学雑誌、査読有、68、2013、189-196、<http://dx.doi.org/10.1265/jjh.68.189>

## 6. 研究組織

### (1) 研究代表者

龍田 希 (TATSUTA, Nozomi)  
 東北大学・医学系研究科・助教  
 研究者番号：40547709

### (2) 連携研究者

仲井 邦彦 (NAKAI, Kunihiko)  
 東北大学・医学系研究科・教授  
 研究者番号：00291336

黒川 修行 (KUROKAWA, Naoyuki)  
 宮城教育大学・教育学部・講師  
 研究者番号：30431505