

百羅漢 ～実験参加者の言語能力差の統制のための漢字テスト

近藤 公久^{*1}
天野 成昭^{*2}

2013年 4月

JCSS-TR-69

日本電信電話株式会社 コミュニケーション科学基礎研究所
〒619-0237 京都府相楽郡精華町光台 2-4
<http://www.kecl.ntt.co.jp/>

Copyright (c) 2013 NTT Corporation.

*1 現在 ATR 知能口ボテイクス研究所 tkondo@atr.jp

*2 現在 愛知淑徳大学

日本認知科学会
事務局
〒464-8601
愛知県名古屋市千種区不老町
名古屋大学 大学院教育発達科学研究科内
電子メール : jcsl@jcsl.gr.jp

百羅漢 ～実験参加者の言語能力差の統制のための漢字テスト

要旨

言語を刺激として実験に用いる研究において刺激の言語特性の統制の重要性が認識されている。そのため、実験刺激セットを作成する場合には頻度や親密度等の特性が明らかな単語群から単語を選択し、実験条件間でバランスをとることは必須である。しかし、実験の参加者については年齢や性別などを大枠統制する以外の特性を考慮することは少なかった。本報告では、実験の参加者の言語的な能力を推定することにより、参加者の差の影響を検討することを目的として作成した百羅漢テストの性質についてまとめた。

はじめに

実験に用いる言語刺激の特性（以後まとめて語特性と呼ぶ）が結果に影響する。たとえば単語の頻度、親密度などは、影響の大きい語特性の代表例であり、出版物等に出現する頻度が高く、よく見かける単語や、親密度が高く、なじみのある単語は処理が速い（e.g., 天野と近藤, 1999a; Connine, Mullennix, Shernoff & Yelen, 1990）。このため、言語心理学的実験、脳科学・神経科学的実験を始めとする、言語刺激を用いるあらゆる実験において、比較対象となる刺激セット間の語特性を統制することは必須である。

一方、実験参加者の言語能力の違いが実験結果に影響を及ぼすことは明らかであるが、これまで子供の発達過程や教育的な学習到達度を直接対象にする研究以外では統計的に分散することを過程して処理し、成人を対象とした実験の参加者の違いによる影響を考慮することは少なかった。しかし、近藤と天野(1998, 2001)は、独自に作成した 100 語の漢字単語の読みのテスト（百羅漢、100 RAKAN: Reading Ability test for KANji words）を実施した結果から、同じ大学の同年代の学生であっても個人差は大きく、言語能力の違いの影響は無視できないことを指摘した。特筆すべきは、大学間差があることも明確であるが、同じ大学内（学部）の学生であっても得点が幅広く分布している点にある（付録 1 参照）。

近年、NTT コミュニケーション科学基礎研究所監修による「NTT データベースシリーズ『日本語の語彙特性』」（天野と近藤, 1999a, 1999b, 1999c; 天野と近藤, 2000; 天野, 笠原, 近藤, 2008; 近藤と天野, 1999a, 1999b, 1999c; 佐久間, 伊集院, 伏見, 辰巳, 田中, 天野, 近藤, 2005）が実験刺激の統制のために用いられるようになっている（以後このデータベースシリーズ全体を指して『NTT データベース』と呼ぶ）。この NTT データベースが収録している語特性の多くは主観的評定値であり、実験参加者とデータベースの各特性の評定者との言語能力の違いについても考慮する必要がある。すなわち、実験参加者と語特性の評定者の間に違いがある場合、その実験刺激の統制に実験参加者とは異なる質の評定者によって評定された語特性を用いていいかという問題が指摘される。幸い、NTT データベース収録の各特性値の評定者は、上述の「百羅漢」の得点が一定の範囲に統制されており、評定者の平均点と分散、および得点の範囲が各データベースの解説の中で公開されている。したがって、「百羅漢」テストを用いれば、実験参加者の得点とデータベース中の特性の評定者の比較が可能であるため、実験参加者のスクリーニングに利用可能と考えられる。

本論文では、まず「百羅漢」テスト項目の特性について解析した結果を示す。次に、幅広い年齢層（中学生から大学生、社会人、高齢者層までを含む）の多人数に対して行った「百羅漢」得点の傾向を分析する。「百羅漢」テスト項目が日本語の単語の様々な特性を軸としてどのような性質を持つ単語から成り立っているか、また、多くの実データによる得点傾向を示すことによって、本テストを使用する際に考慮すべき点を明らかにする。

百羅漢

概要

「百羅漢」は NTT データベース収録の各語特性の評定実験参加者を統制するために作成・実施されたスクリーニングテストである。このテストでは、日本語漢字単語 100 語の読みを記述し、その正答数を得点とする。百羅漢のテスト項目として選択された単語は、(1) 漢字のみで表記される単語、かつ、(2) 同字異音語、同音異字語が存在しない単語、かつ、(3) 文字提示における単語親密度が 4.0 以下の単語、かつ、(4) 音声提示における単語親密度が 4.5 以上の単語である(近藤と天野, 1998)。つまり、

「百羅漢」のテスト項目は、文字（漢字）で表記された単語を見たときには親密度が低い（なじみがない）単語であるが、音声で聞いたときには親密度が高い（なじみがある）単語であることを想定している。

「百羅漢」のテスト項目を表1に示した。テスト項目のIDは、文字提示における単語親密度の高い単語から低い単語の順となっており、印刷されたテスト用紙もこの順に並べられている。したがって、受験者は平均的に易しい単語から難しい単語の順に回答を求められることになる。表中、正答率はその項目の正答率(%)[N=1410]、困難度2と識別力2はそれぞれ2母数ロジスティックモデルにおける困難度と識別力の推定値、困難度1は1母数ロジスティックモデルにおける困難度の推定値を示している。

百羅漢テスト項目の単語特性

本節では、百羅漢テスト項目として出現する単語および文字の特性を解析する。但し、百羅漢テスト作成時に用いた親密度値は、Amano, Kondo, and Kakehi (1995) (以後AK1995と呼ぶ)の解析に用いた親密度値であり、NTTデータベースの単語親密度(天野と近藤, 1999a)とは異なる。そこで、単語親密度については、AK1995とNTTデータベースの単語親密度の両方について解析を行った。しかし、NTTデータベースでは、表記妥当性の低い表記に対しては文字提示による単語親密度の評定がなされていないため、文字提示時の単語親密度が得られない単語が16項目存在した。このため、単語の表記に関わる特性についてはこれらの16項目を除いて解析した。また、表1の代替漢字欄に示したとおり、これら16項目は、漢字表記が難しく「漢字かな交じり」もしくは「ひらがな」表記で親密度評定がなされたもの(例: 目脂→目やに、水馬→あめんぼ) [これらは表中「一」で示した]、異なる漢字(異体字)が用いられたもの(例: 貫祿→貫禄)、繰り返しを「々」で表記したもの(例: 茶茶→茶々)である。なお、「万万」はNTTデータベースにもこのままの表記で親密度データが存在するため解析から除外しなかった。

一 単語親密度

単語親密度は、天野と近藤(1999a)は「なじみの程度」と総称している主観的評定値であり、評定者32名による1(なじみがない)から7(なじみがある)の7段階評定によって得られた値である。同書の表1-5に示される評定基準のアンケート結果から、「単語親密度は、使用頻度、遭遇頻度、意味理解度、既知度等と関連が深い」とされている。これらに加え、文字単語親密度(文字で表記された単語に対する評定値)では、「可読性」、つまり、文字で表記された単語の読みが容易かどうかも関わっている。百羅漢のテスト項目選択においては、親密度を基準に用いている。そこでまず、テスト項目の単語親密度の分布を確認する。特に、百羅漢テスト作成時に用いた親密度値は天野と近藤(1999a)収録の値とは異なるため、両者の比較も行う。

なお、単語親密度評定者40名の百羅漢テスト得点の平均値は72.1、標準偏差は6.71、最大値は90、最小値は61である(天野と近藤, 1999a)。

図1a, 1bに示されるように、AK1995の単語親密度による選択基準から、音声提示時の単語親密度(afam)は高く(afam \geq 4.5)、文字提示時の単語親密度(vfam)は低い(vfam \leq 4.0)ことが確認できる。つまり、百羅漢テスト項目のAK1995による音声および文字単語親密度の分布はこれらの基準で切断されている。また、この基準を満たす範囲では元のデータベース全体(新明解国語辞典第二版)の各単語親密度分布(N=62,001)と類似した分布傾向となっている(各図の折れ線)。

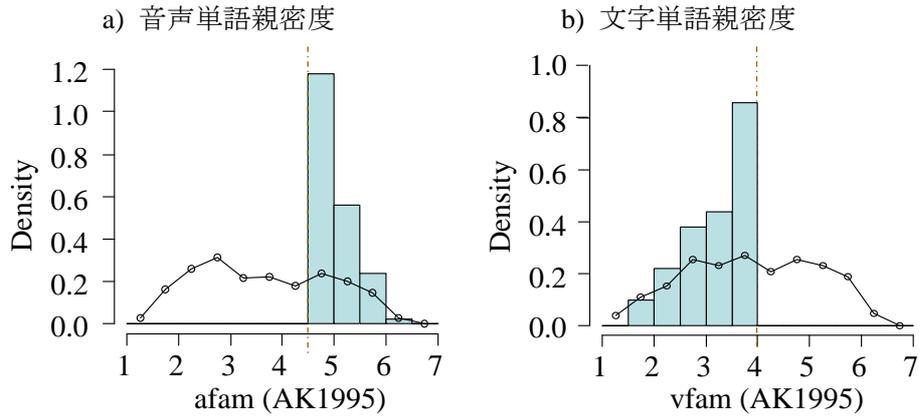


図1 テスト項目の単語親密度(AK1995)の頻度分布

一方、テスト項目に対する NTT データベースの単語親密度の音声提示による単語親密度（音声単語親密度: afam）、文字提示による単語親密度（文字単語親密度: vfam）、および、音声と文字同時提示による単語親密度（音声文字単語親密度: avfam）の分布を図 2a,2b,2c に示す（但し、vfam と avfam については N=84）。各図の折れ線は NTT データベース全体の分布を示している。

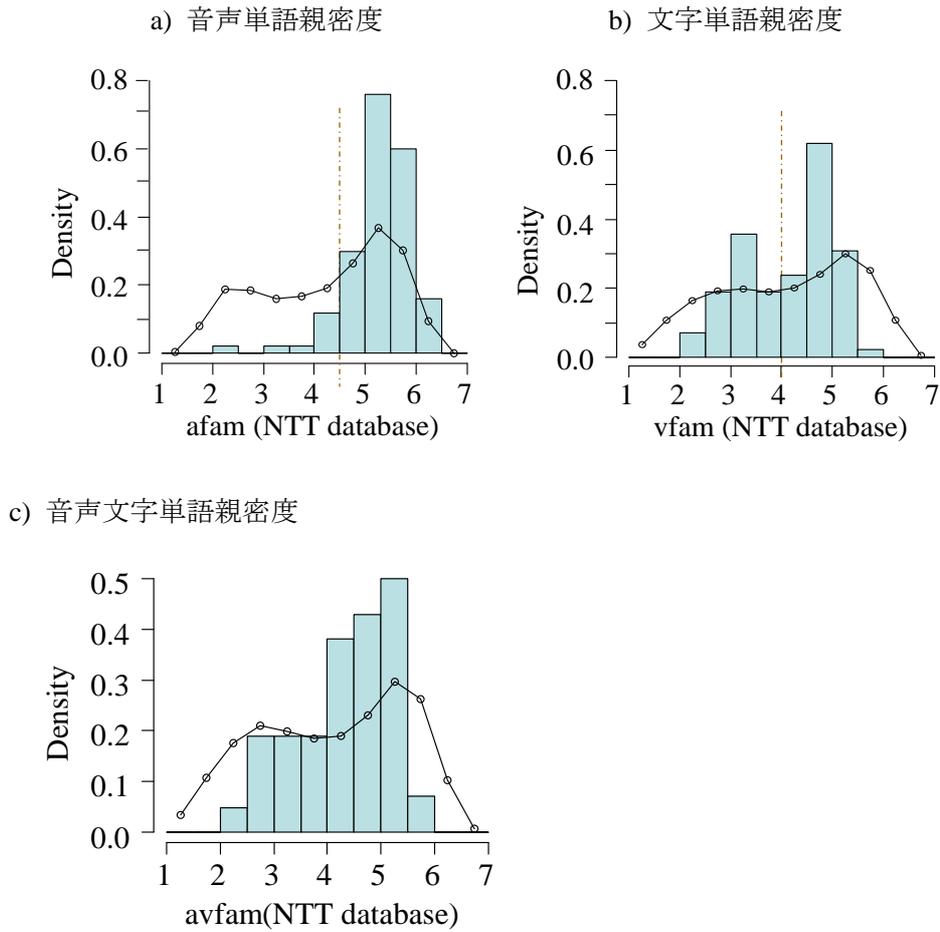


図2 テスト項目の単語親密度の頻度分布

図 2a からわかるように、AK1995 による選択基準を反映して、NTT データベースによる afam でみてもテスト語には低親密度語がほとんど存在しない。これは、同図のデータベース全体(N=69,084)の afam の分布 (折れ線) と大きく異なっている。ちなみに、afam が 4.5 未満となった単語は、唐物 (とうぶつ)、天朝 (てんちょう)、除数 (じょすう)、万万 (ばんばん)、参画 (さんかく)、地平 (ちへい)、干菓子 (ひがし)、奇聞 (きぶん)、交直 (こうちよく) であった。

これに対し、vfam (図 2b) は、極端に低い親密度および極端に高い親密度を有する語が存在しないことを除くと、データベース全体(N=69,084)の vfam の分布と類似している。この傾向は avfam も同じである (図 2c)。したがって、百羅漢テスト項目の選択において選択基準とした、音声提示において単語親密度が高く、文字提示において単語親密度が低いという条件は、NTT データベースの親密度値でみると前者のみしか実現されていないが、音声で聞けば単語が容易に理解できる単語を対象としてその漢字表記の読みをテストするという基本概念は NTT データベースの単語親密度を基準としてみた場合でも実現していると考えてよい。このことは、百羅漢作成の当初のもくろみを必ずしも実現していないが、幅広い受験者の能力を反映するという意味ではかえって有効な場合もあると期待できる。

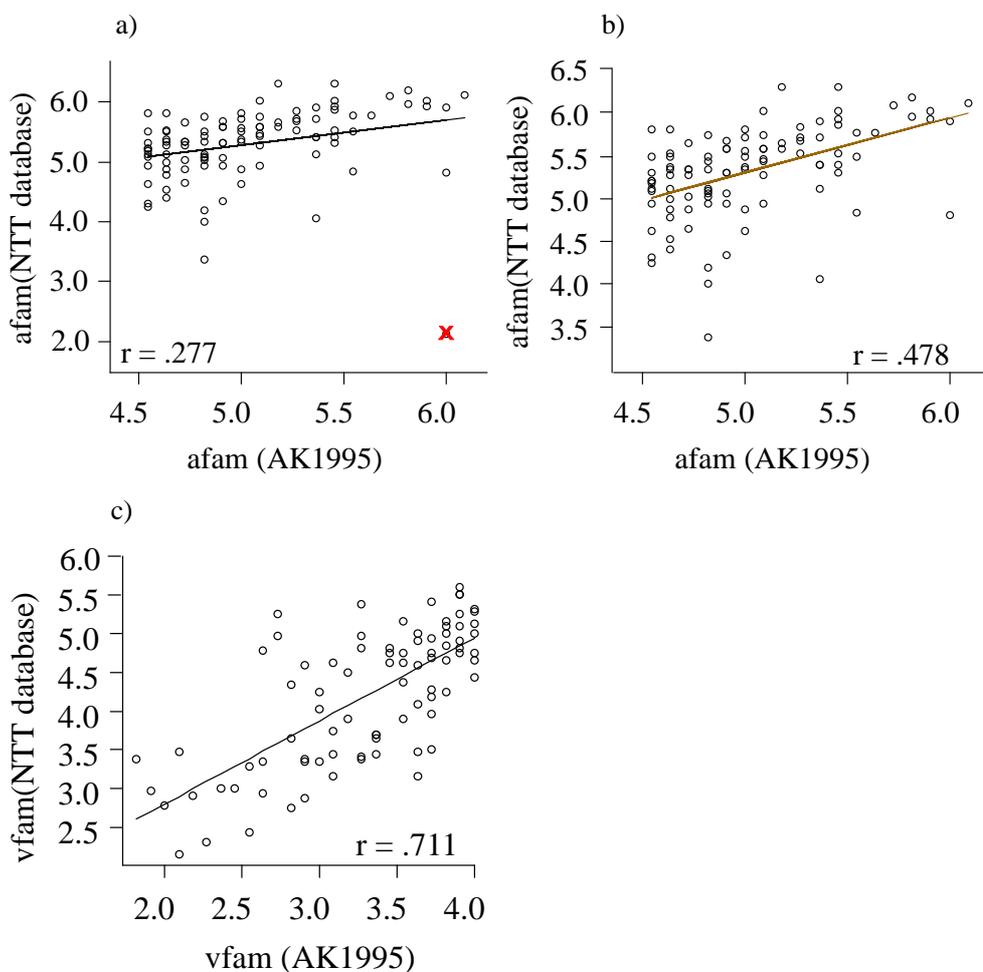


図 3 NTT データベースシリーズの単語親密度と AK1995 の関係

図 3 に、百羅漢テスト項目の AK1995 による単語親密度と NTT データベースによる単語親密度の相関図を示す。この両者の相関は統計的に有意ではあるが相関係数は小さい[$r = .277, p < .01$]。図 3a において 1 例 (右下×印) が悪影響を及ぼしている可能性を指摘できる。その単語は、「唐物 (とうぶつ)」である。そこで、この 1 語を除いた相関係数を求めると、 $r = .478$ となり中程度の相関がみられた (図 3b)。また、vfam では、 $r = .711 (p < 0.001, N=84)$ と高い相関がみられた (図 3c)。

以上から、AK1995 を用いて作成された百羅漢のテスト項目の単語親密度は、AK1995 と NTT データベースの親密度に対応に幅がある (中程度の相関) 分、百羅漢テスト項目の選定基準とした音声単語親密度が高いもののみを選択すると基準に若干の幅ができていと考えられる。また文字単語親密度においては、両者に高い相関がみられているが、AK1995 の文字単語親密度は NTT データベースの文字単語親密度と比較して低めにシフトしているため、AK1995 の文字単語親密度が低い単語のみを選択したにもかかわらず、NTT データベースの文字単語親密度では親密度が高いテスト項目が含まれる結果となっていることがわかる。

一表記妥当性

日本語の単語は、漢字かな交じり、ひらがな、カタカナで表記可能であるが、それぞれの単語がどのような表記で日常用いられるかは、おおよそ決まっている。NTT データベースシリーズの表記妥当性は、「ある単語を表記するのにどんな表記がもっともらしいか、逆に、どんな表記だともっともらしさに欠けるかを示す指標」であり、評定者 20 名による 1 (もっともらしくない) ~ 5 (もっともらしい) の 5 段階評定値である (近藤と天野, 1999a)。また、表記妥当性は、浮田ら (浮田, 杉島, 皆川, 井上, 賀集, 1996) の主観的表記頻度と相関が高いことが示されている (近藤と天野, 1999a)。

なお、表記妥当性評定者 20 名の百羅漢テスト得点の平均値は 69.0、標準偏差は 6.8、最大値は 86、最小値は 60 である (近藤と天野, 1999a)。

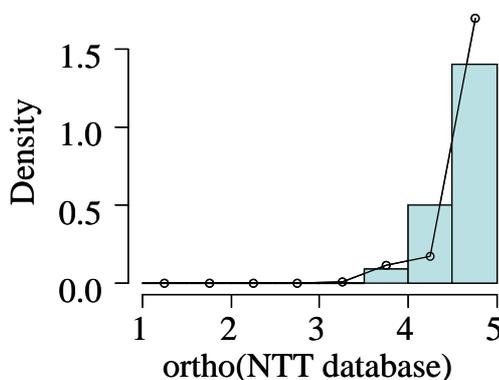


図 4 表記妥当性の分布

図 4 に百羅漢テスト項目の表記妥当性(ortho)の分布 (棒) と NTT データベース全体の分布 (折れ線) を示す。この図から百羅漢のテスト項目の表記妥当性は、データベース全体の分布と類似しているものの、若干低めの単語が多い傾向にあることがわかる。これは、音声親密度を一定以上高くし、文字親密度は低い単語を選択基準としたことを反映した結果と考えられる。ちなみに、上述した表記が NTT データベースの単語親密度値が存在しない 16 項目を除いて表記妥当性が 4.0 未満のテスト項目は、「御日様」、「笹」、「万万」の 3 項目であった。また、百羅漢テスト項目に用いた表記が各単語の表記妥当性が最大とならない項目は、「金蔓」、「薺医者」、「髭面」、

「湯煙」、「誰彼」、「鰻井」の6項目で、それぞれ、「金づる」、「ヤブ医者」、「ヒゲ面」、「湯けむり」、「誰かれ」、「うな井」が最大表記妥当性を有する表記であった。

－単語出現頻度

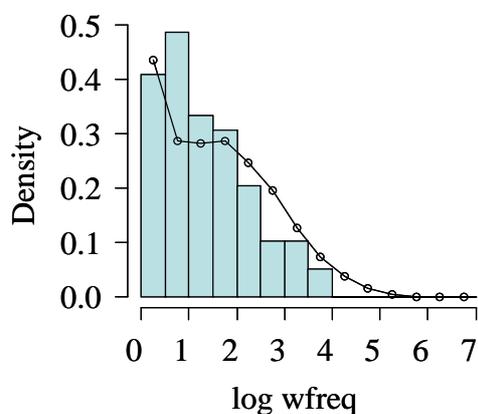


図5 単語出現頻度の分布

図5に百羅漢テスト項目の単語出現頻度(log wfreq)の分布(棒)とNTTデータベース全体の分布(折れ線)を示す。図5から、百羅漢テスト項目の単語出現頻度は、10回から100回程度が中心で、1000回を超える単語は非常に少ない。このことは、NTTデータベース全体の単語出現頻度分布と比較して、出現頻度が少ない単語に偏っていることがわかる(漢字単語の平均は2626回)。この結果も、音声単語親密度が一定以上の単語に制限したセットであっても、文字単語親密度が低いものを多く含むセットであることを反映していると言える。

百羅漢テスト項目の漢字特性

－漢字親密度(語中最小値)(NTTデータベース)

図6に百羅漢テスト項目に含まれる各語中の最低漢字親密度(kfam(min))の分布(棒)およびNTTデータベース全体に出現する単語中の最低漢字親密度の分布(折れ線)を示す。図6からわかるように、漢字の親密度(語中最小値)は、データベース全体の分布と比較して、3未満の漢字の出現が少ないことを示している。

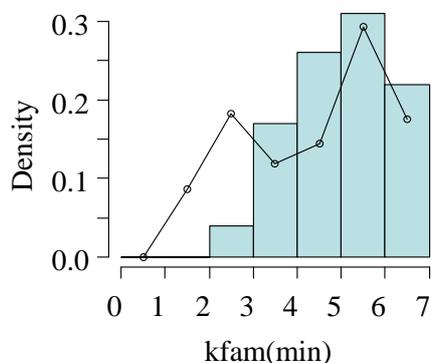


図6 漢字親密度(語内最小値)の分布

したがって、本テスト項目では極端に漢字親密度が低い漢字は用いられていないといえる。つまり、知らない漢字が出現することによる影響は比較的小さく、単純な漢字の知識を問うテストとは異なる指標となりうると考えられる。

なお、漢字親密度評定者 20 名の百羅漢テスト得点の平均値は 72.8, 標準偏差は 8.4, 最大値は 90, 最小値は 60 である(近藤と天野,1999c)。

－漢字頻度（語中最小値）(NTT データベース)

図 7 に、百羅漢テスト項目にある語中漢字頻度の最小値 (kfreq(min)) (棒) と NTT データベース全体に存在する語中漢字頻度の最小値 (折れ線) を示す。図 7 から、語中漢字頻度の最小値は百羅漢テスト項目とデータベース全体で同様の分布をしている。したがって、本テスト項目に用いられている漢字は新聞中と比較して極端に難しい漢字が多い訳ではないと考えられる。このことは、漢字親密度による検討結果と同じ結果である。

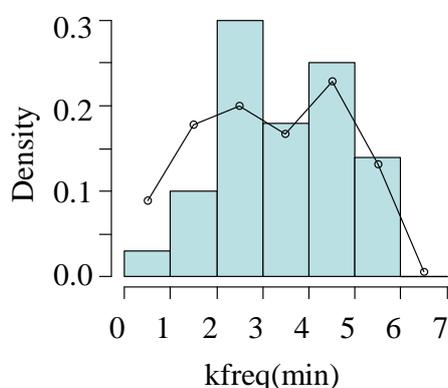


図 7 漢字頻度（語内最小値）の分布

－学年配当（語中最大値）(NTT データベース)

図 8 に、百羅漢テスト項目にある語中漢字の最大学年配当 (gakunen (max)) (棒) と NTT データベース全体に存在する語中漢字の最大学年配当 (折れ線) を示す。図 8 からわかるように、学年配当（語中最大値）は、データベース全体と同様の分布をしている。

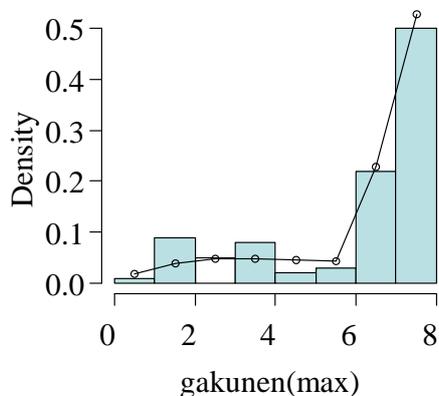


図 8 学年配当（語内最大値）の分布

したがって、本テスト項目に用いられている漢字は極端に学年配当が高い訳ではないと考えられる。このことは、漢字親密度および漢字出現頻度による検討結果と同じ結果である。

一画数（語中最大値）

図9に、百羅漢テスト項目にある語中漢字の最大画数（stroke(max)）（棒）とNTTデータベース全体に存在する語中漢字の最大画数（折れ線）を示す。図9からわかるように、画数（語中最大値）は、データベース全体と同様の分布をしている。したがって、本テスト項目に用いられている漢字は極端に画数が多い漢字が多い訳ではないと考えられる。このことは、漢字親密度、漢字出現頻度、学年配当による検討結果と同じ結果である。

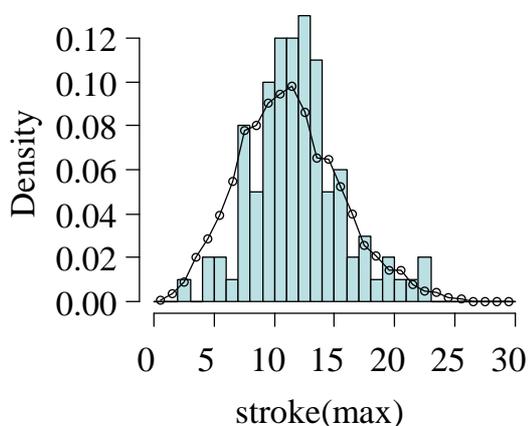


図9 画数（語内最大値）の分布

一読みの妥当性

図10aに百羅漢テスト項目にある語中漢字の読みの妥当性平均（yplaus(avg)）、図10bに百羅漢テスト項目の先頭漢字の読みの妥当性（yplaus(first)）を示す。それぞれ、棒がテスト項目、折れ線がNTTデータベース全体の単語を対象にした値を示す。

a) 語中漢字の読みの妥当性平均

b) 先頭漢字の読みの妥当性

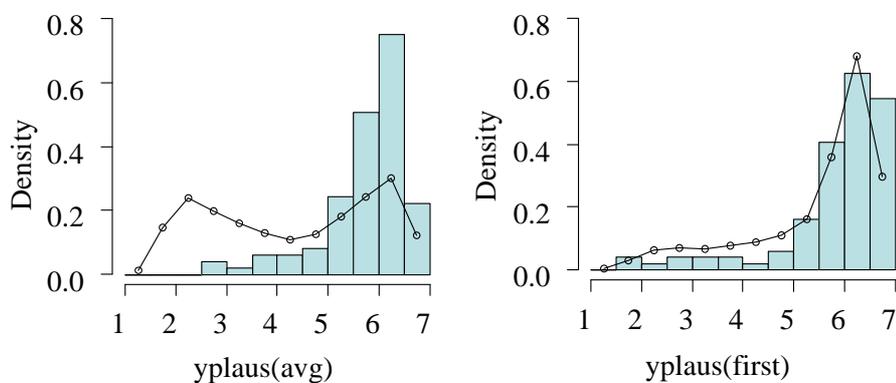


図10 読みの妥当性の分布

図 10 からわかるように、読みの妥当性は、最初の漢字では NTT データベース全体と同様の分布をしている。しかし、平均では多少異なる。これについては後で考察する。

なお、読みの妥当性評定者 24 名の百羅漢テスト得点の平均値は 72.9、標準偏差は 5.8、最大値は 83、最小値は 63 である(近藤と天野,1999c)。

考察

百羅漢のテスト項目は、上述のとおり、総じて NTT データベースによる特性について特に偏りのある特別な分布を示すものではない。テスト項目選択において音声単語親密度が高く、文字単語親密度が低い単語を条件とした。この特性を保持しつつ、NTT データベース中に存在する幅広い範囲から選択されているといえる。

「百羅漢」得点データ収集実験

目的

百羅漢を作成した目的は、NTT データベース収録の各語特性の評定実験参加者を統制するためであった。それぞれの評定実験に参加した参加者の百羅漢得点データは、前章のそれぞれの特性値の解析中に示している。

ここでは、百羅漢得点の一般的傾向を把握するために実施した実験について示す。幅広い年代層の得点傾向を把握することで、NTT データベースの各特性の評定者が本テストの得点でどのようなレベルであったかを知ることができ、また、本テストの適用対象を検討する上でも有効であると考ええる。

実験方法

実験参加者： 実験参加者は、中学生 49 名、高校生 56 名、予備校生 5 名、専門学校生 31 名、短大生 7 名、大学生 226 名、大学院生 14 名、社会人 653 名、合計 1041 名であった。男女比はほぼ半々であった（男性 522 名、女性 519 名）。また、年齢は 12 歳から 76 歳であった。なお、実験への参加に際しては、実験の目的および取得データの扱いについての説明を事前に行い、同意書へ署名することによって自由意志による参加の意思の確認を行った。さらに、中高生の参加に際しては、保護者にも同様の説明と同意書への署名による実験参加承諾確認を行った。実験参加者には所定の謝金を支払った。

手続き： 表 1 のテスト項目を ID 順に印刷した用紙に、各語の読みをひらがなで記入してもらった（付録 2 参照）。所要時間は 10 分とした。実際の実験は、他のアンケート調査と合わせて行い、全体で 1 時間程度、説明を含めて約 1 時間半程度を要した。なお、アンケート調査の詳細は、小林、天野、正高(2007)に示されている。

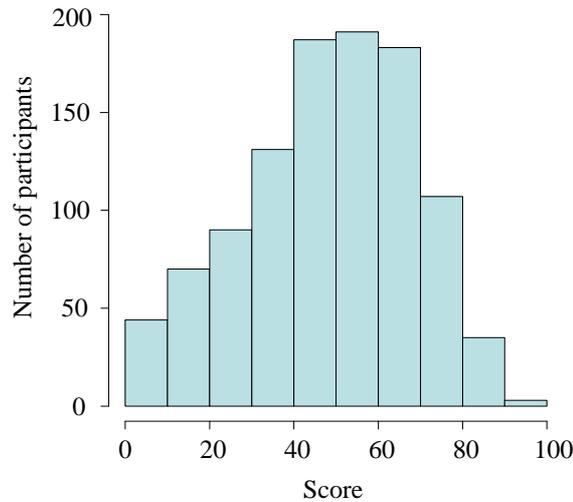


図 11 百羅漢得点分布 (N=1041)

結果

本実験参加者には、中学生や高校性を含んでおり、異なる分布を示すと予想される。そこで、参加者の群別による検討を行った。図 12 に参加者の群別の分布と NTT データベースの親密度評定者の得点分布を示す。図 12 から得点の低い範囲の多くは中高生であることがわかる。

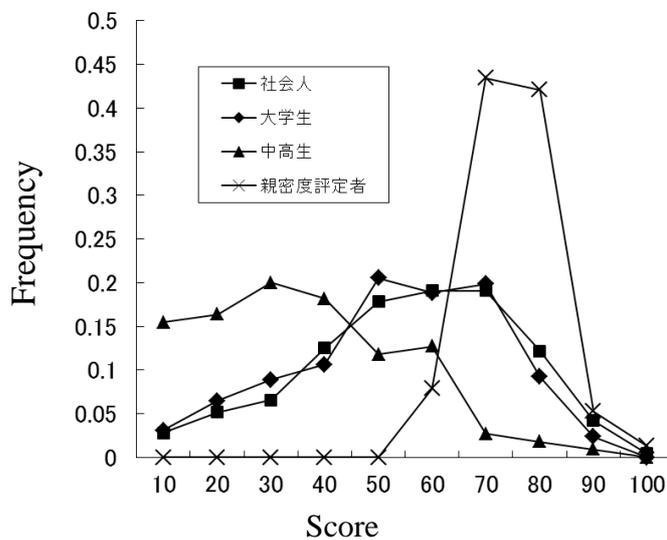


図 12 群別分布と親密度評定者との比較

一方、社会人と大学生および専修学校生の百羅漢得点分布はほぼ一致している。しかし、社会人と大学生および専修学校生の百羅漢得点分布と NTT データベースの単語親密度評定実験参加者の百羅漢得点分布は大幅に異なる。これは、NTT データベース構築のために募集した単語親密度評定実験参加者は、前述のとおり百羅漢得点が 60 点以上に限られているためである[N=40; mean=72.1, SD=6.71, Max=90, Min=61]。このことは、NTT データベースの単語親密度を用いて実験刺激を統制した実験を実施する

場合には、実施する実験の参加者が社会人もしくは大学生および専修学校生であっても、参加者の百羅漢得点分布がこの NTT データベースの単語親密度評定実験参加者の分布と大幅に異なる可能性があることを示唆する。

但し、それぞれの実験の参加者の百羅漢得点分布が NTT データベースの単語親密度評定実験参加者の分布と一致しなければならないわけではない。しかし、単に大学生であること、あるいは、〇〇大学の学生であることだけでは、参加者の言語的能力には幅があるため十分な検討が必要であること、NTT データベースの単語親密度を用いて統制した言語刺激が、それぞれの実験の目的どおりの統制となっているかについても検討が必要であることは明らかである。

なお、「日本語の語彙特性」データベースの他の特性値の評定実験参加者にも実施している。それぞれの平均点数と分散は大幅には異なるないので、各特性値を用いて統制した実験を実施する場合には単語親密度と同様の注意が必要である。

百羅漢テスト項目の分析

本論文では、百羅漢のテスト項目の単語特性および漢字の特性を示した。この結果、百羅漢テスト項目の選択時の選択基準であった、音声単語親密度は高く、文字単語親密度がある程度低い単語、であることを除き、百羅漢テスト項目の特性値の分布は NTT データベース全体の各特性値分布と大きな違いがないことが示された。つまり、本テスト項目は NTT データベース全体からバランスよく選択されているといえる。

本章では、実験によって得られた百羅漢の各テスト項目の正誤データから、各テスト項目と百羅漢合計得点の関係について考察した結果を示す。

項目分析

項目反応理論を用い、各項目がテストにおいてどのように貢献しているかを確認する。表 1 に 2 母数ロジスティックモデルによる困難度(困難度 2)と識別力(識別力 2)、および、1 母数ロジスティックモデルによる困難度(困難度 1)を示している。図 13(a) に 1 母数による項目反応曲線、図 13(b) に 2 母数による項目反応曲線を示す。

表の値および図 13(a)は、ability が高くても得点できない項目がかなりあることを示している[X47: 彼是, X59: 水馬, X70: 惚気, X82: 態態, X89: 策 X07: 空便, X52: 唐物, X80: 正舷, X85: 遙遙]。また、図 13(b)は、ability の上昇によって正答率が上昇しない識別力が低いものがあることを示している[X07: 空便, X52: 唐物, X80: 正舷, X85: 遙遙]。この結果は、少数項目でのテスト（たとえば 20 語）を作成する際の参考になる。たとえば、幅広い受験者の差を見たい場合には、識別力の高い項目から困難度のある程度幅広くとり、ある程度以上の高得点が予想される受験者間の差を検討したい場合には困難度の高いものを選択するなどが考えられる。少数項目でのテストの構築については今後の課題とする。

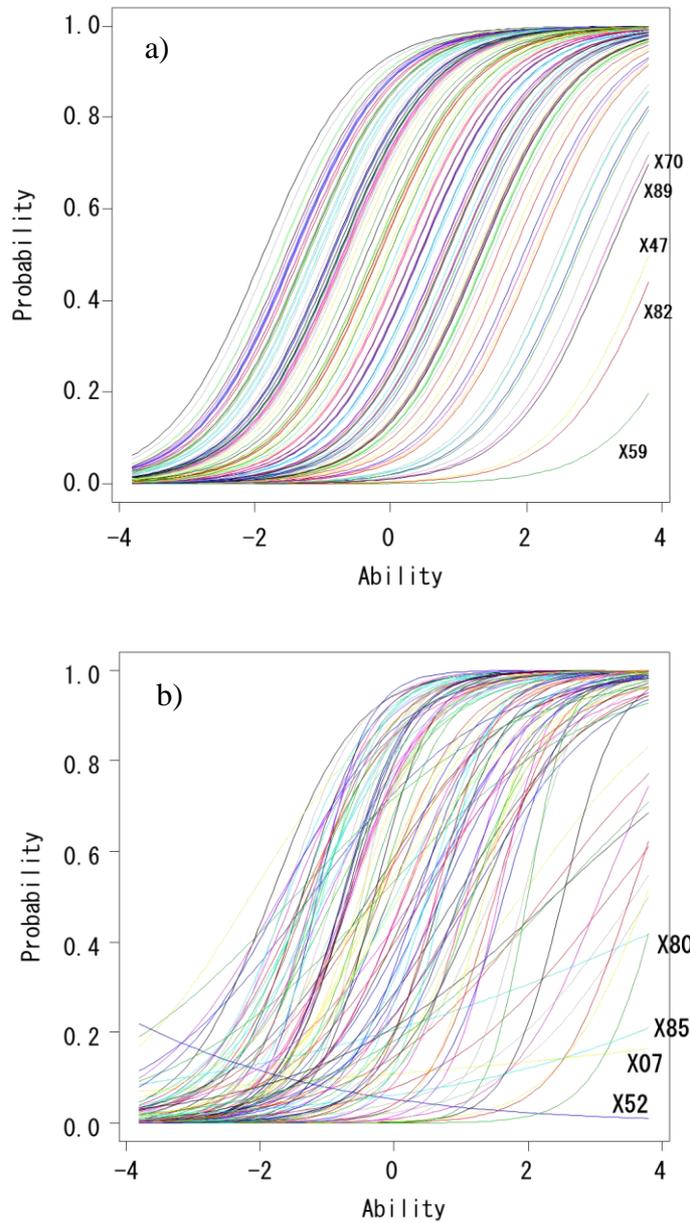


図 13 項目反応曲線

a) 1 母数ロジスティックモデル、b) 2 母数ロジスティックモデルによる推定結果

個人の特性(Ability)を θ とするとき、項目 j を正答する正答確率(Probability)を $P_j(\theta)$ 、識別力を a_j 、困難度を b_j 、と表すとき、それぞれのモデルは、以下で表される。

$$P_j(\theta) = 1 / (1 + \exp(-D(\theta - b_j))) \quad (a)$$

$$P_j(\theta) = 1 / (1 + \exp(-Da_j(\theta - b_j))) \quad (b)$$

但し、尺度化係数 $D=1$ とした。

まとめ

本論文では、百羅漢得点分布が社会人や大学生においても幅広く分布すること、その分布が、NTT データベースの各特性値の評定実験参加者の得点分布と異なることを示し、単語を刺激として用いる実験において、参加者の言語能力を吟味する必要性を示した。百羅漢得点を指標とすることの妥当性については、既に、近藤と天野(2001, 2006)、文章の読み (Kondo, Wydell, and Amano, 1998)でその効果を示している。また、高齢者の聴力との関係(近藤, 伊集院, 天野, 2008)や知能テスト(日本版 WAIS-III)の言語得点との関係も検討されている。今後、さらに多くのデータを集め、様々な言語心理学的データと語特性と参加者の言語能力(百羅漢得点)の関係を明らかにする必要がある。

文献

- Amano, Kondo, and Kakehi (1995) Modality dependency of familiarity ratings of Japanese words, *P&P*, 57(5), 598-603.
- 天野成昭,近藤公久 (1999a)."日本語の語彙特性 第1巻 単語親密度",NTTデータベースシリーズ,三省堂.
- 天野成昭,近藤公久 (1999b)."日本語の語彙特性 第3巻 アクセント",NTTデータベースシリーズ,三省堂.
- 天野成昭,近藤公久 (1999c)."日本語の語彙特性 第4巻 品詞",NTTデータベースシリーズ,三省堂.
- 天野成昭,近藤公久 (2000)."日本語の語彙特性 第7巻 頻度",NTTデータベースシリーズ,三省堂.
- 天野成昭,笠原要,近藤公久 (2008)."日本語の語彙特性 第9巻 単語親密度補遺",NTTデータベースシリーズ,三省堂.
- Connine, Cynthia M.; Mullennix, John; Shernoff, Eve; Yelen, Jennifer, Word familiarity and frequency in visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, Vol 16(6), Nov 1990, 1084-1096.
- 金田一京助, 金田一春彦, 見坊豪紀, 柴田武, 山田忠雄 (編) (1974) 新明解国語辞典第二版, 三省堂.
- 金田一京助, 柴田武, 山田明雄, 山田忠雄 (編) (1989) 新明解国語辞典第四版, 三省堂.
- 小林哲生, 天野成昭, 正高信男 (2007) モバイル社会の現状と行方,NTT出版
- Kondo, Wydell, and Amano (1998) *Language Processing and Reading Skills in Japanese*. CUNY conference on human sentence processing, New Brunswick, NJ.
- 近藤公久,天野成昭 (1998)."漢字単語の読み能力テスト: 単語親密度を利用した言語能力の推定法" 日本心理学会第62回大会.
- 近藤公久,天野成昭 (1999a)."日本語の語彙特性 第2巻 単語表記",NTTデータベースシリーズ,三省堂.
- 近藤公久,天野成昭 (1999b)."日本語の語彙特性 第5巻 漢字",NTTデータベースシリーズ,三省堂.
- 近藤公久,天野成昭 (1999c)."日本語の語彙特性 第6巻 漢字",NTTデータベースシリーズ,三省堂.
- 近藤公久,天野成昭 (2001)."漢字単語の読み能力テスト「百羅漢」の得点傾向" 日本心理学会第65回大会.
- 近藤公久,天野成昭 (2006). 漢字単語の読み能力テスト「百羅漢」の項目分析 日本心理学会第70回大会.
- 近藤公久,伊集院睦雄,天野成昭 (2008). 百羅漢の得点から見た言語能力の加齢変化, 日本心理学会大会.
- 日本版 WAIS-III 刊行委員会訳編 (2006) 日本版 wais-III 日本文化科学社 [Wechsler, D. (1997) Harcourt Assessment Inc.].
- 佐久間尚子, 伊集院睦雄, 伏見貴夫, 辰巳格, 田中正之, 天野成昭, 近藤公久 (2005)."日本語の語彙特性 第8巻 単語親密度",NTTデータベースシリーズ,三省堂.
- 浮田潤, 杉島一郎, 皆川直凡, 井上道雄, 賀集寛 (1996) , “日本語の表記形態に関する心理学的研究”, 心理学モノグラフ, 25.

表1 百羅漢テスト項目と正答率および項目分析結果

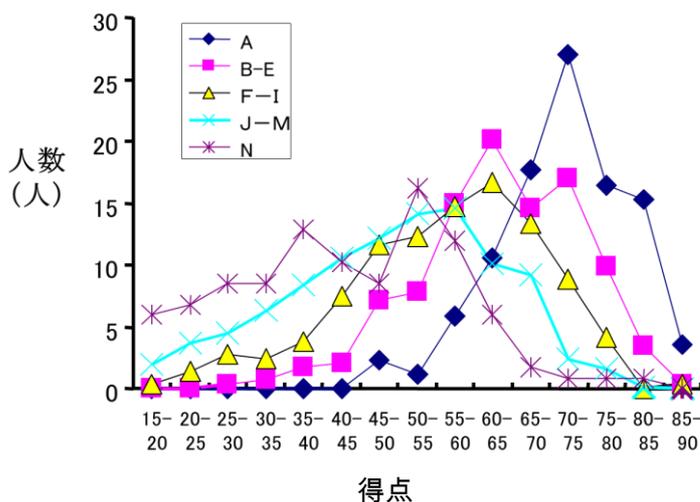
No	項目	正答	代替漢字	正答率	困難度2	識別力2	困難度1
1	狸	タヌキ		88.9	-1.844	1.539	-1.886
2	幹旋	アッセン		58.1	-0.200	1.577	-0.170
3	貫祿	カンロク	祿	82.1	-1.184	2.036	-1.364
4	湯煙	ユケムリ		71.7	-1.252	0.795	-0.774
5	悠悠	ユウユウ	々	83.6	-1.390	1.651	-1.462
6	湾曲	ワンキョク		84.1	-1.509	1.483	-1.503
7	空便	カラビン		11.1	16.849	0.124	2.089
8	無恥	ムチ		69.2	-0.642	1.658	-0.654
9	巡業	ジュンギョウ		84.7	-1.375	1.902	-1.545
10	餌食	エジキ		69.7	-0.664	1.675	-0.681
11	民放	ミンポウ		80.2	-1.303	1.391	-1.242
12	庭草	ニワクサ		40.2	0.519	1.358	0.553
13	薄着	ウスギ		86.0	-1.590	1.576	-1.640
14	出土	シュツド		73.9	-0.985	1.302	-0.886
15	曲者	クセモノ		58.4	-0.242	1.351	-0.182
16	勤続	キンゾク		87.6	-1.615	1.789	-1.774
17	滑稽	コッケイ		72.5	-0.706	2.117	-0.817
18	奴等	ヤツラ		80.3	-1.365	1.296	-1.247
19	御日様	オヒサマ		71.5	-0.889	1.236	-0.765
20	亡者	モウジャ		43.0	0.412	1.231	0.436
21	除数	ジョスウ		73.6	-1.032	1.184	-0.871
22	夜長	ヨナガ		70.2	-0.660	1.798	-0.704
23	地平	チヘイ		85.8	-2.149	0.960	-1.625
24	執拗	シツヨウ		68.4	-0.475	2.919	-0.618
25	蓋	フタ		71.5	-0.679	1.997	-0.765
26	滞日	タイニチ		18.6	2.241	0.783	1.562
27	火蓋	ヒブタ		60.1	-0.244	1.919	-0.254
28	老衰	ロウスイ		83.4	-1.146	2.666	-1.449
29	夜色	ヤシヨク		21.4	5.046	0.267	1.404
30	鎧	ヨロイ		82.1	-1.707	1.050	-1.364
31	改悪	カイアク		59.8	-0.297	1.392	-0.242
32	髭	ヒゲ		74.9	-0.776	2.367	-0.941
33	輔佐	ホサ		59.2	-0.365	0.977	-0.214
34	小身	ショウシン		58.0	-0.296	1.008	-0.166
35	襟首	エリクビ		79.8	-1.113	1.819	-1.218
36	収賄	シュウワイ		54.6	-0.065	1.529	-0.025
37	俄然	ガゼン		78.3	-1.147	1.479	-1.127
38	満帆	マンパン		36.2	0.652	1.547	0.715
39	茶茶	チャチャ	々	55.7	-0.237	0.791	-0.071
40	磯巾着	イソギンチャク		78.7	-1.158	1.503	-1.149
41	髭面	ヒゲヅラ		62.5	-0.296	2.399	-0.356
42	陰険	インケン		81.6	-1.291	1.586	-1.326
43	干菓子	ヒガシ		28.8	1.026	1.355	1.037
44	贅肉	ゼイニク		73.5	-0.760	2.033	-0.866
45	急須	キュウス		80.6	-1.071	2.182	-1.265
46	酒粕	サケカス		70.5	-0.696	1.682	-0.718
47	彼是	アレコレ	-	1.4	3.758	1.473	3.801
48	天辺	テッペン		4.2	3.794	0.974	2.947
49	腋毛	ワキゲ		33.7	0.695	1.969	0.821
50	奇聞	キブン		30.3	1.078	1.083	0.972

表1(続) 百羅漢テスト項目と正答率および項目分析結果

No	項目	正答	代替漢字	正答率	困難度2	識別力2	困難度1
51	交直	コウチョク		71.182	-1.497	0.625	-0.751
52	唐物	トウブツ		5.668	-6.795	-0.422	2.700
53	鯉	カツオ		77.329	-1.259	1.175	-1.072
54	藪医者	ヤブイシャ		83.189	-1.219	2.146	-1.435
55	細則	サイソク		54.467	-0.073	1.375	-0.021
56	肉饅頭	ニクマンジュウ		76.753	-0.917	1.989	-1.040
57	覇者	ハシャ	覇	45.917	0.279	1.340	0.321
58	塵紙	チリガミ	-	48.895	0.156	1.483	0.202
59	水馬	アメンボ	-	0.384	3.953	2.119	4.784
60	参画	サンカク		43.324	0.355	1.854	0.424
61	誰彼	ダレカレ		40.250	0.490	1.554	0.548
62	辻褄	ツジツマ		49.472	0.133	1.548	0.179
63	筏	イカダ		24.304	1.162	1.589	1.254
64	鬼瓦	オニガワラ		69.549	-0.720	1.424	-0.672
65	弓馬	キュウバ		34.102	0.926	0.982	0.804
66	万万	バンバン		10.663	3.213	0.762	2.131
67	鳴戸	ナルト		54.563	-0.221	0.635	-0.025
68	蹄	ヒヅメ		36.023	0.703	1.329	0.723
69	梯子	ハシゴ		31.508	0.717	2.729	0.917
70	惚気	ノロケ		3.362	3.076	1.483	3.135
71	神州	シンシュウ		21.614	1.888	0.835	1.394
72	粗方	アラカタ		26.129	1.144	1.390	1.164
73	竹藪	タケヤブ		73.967	-0.774	2.075	-0.891
74	天朝	テンチョウ		43.612	0.429	0.968	0.413
75	頓知	トンチ		64.553	-0.413	1.897	-0.444
76	鯖	サバ		82.997	-1.674	1.144	-1.422
77	算盤	ソロバン	-	52.161	0.012	1.204	0.072
78	鰻井	ウナドン		46.206	0.313	0.849	0.309
79	絆創膏	バンソウコウ		66.763	-0.505	1.864	-0.542
80	正絃	ショウゲン		6.820	3.571	0.851	2.538
81	付会	フカイ		23.727	2.378	0.550	1.283
82	態態	ワザワザ	々	1.153	3.521	1.801	3.964
83	十念	ジュウネン		21.134	2.385	0.630	1.420
84	汗疹	アセモ		29.875	1.146	1.009	0.990
85	遙遙	ハルバル	々	6.916	7.407	0.365	2.527
86	青痣	アオアザ		23.343	1.079	2.154	1.303
87	満配	マンパイ	-杯	50.624	0.074	1.044	0.133
88	金蔓	カネヅル		7.589	1.930	2.450	2.444
89	筧	ザル		3.074	2.515	2.448	3.206
90	目脂	メヤニ	-	15.082	1.497	2.080	1.788
91	誰其	ダレソレ	-某	23.055	1.168	1.775	1.318
92	火燧	コタツ	-	12.872	1.637	2.062	1.947
93	黙禱	モクトウ		39.289	0.488	2.063	0.588
94	黴	カビ		12.008	1.561	2.657	2.017
95	椋鳥	ムクドリ		16.427	1.335	2.544	1.698
96	舶載	ハクサイ		20.173	1.450	1.414	1.473
97	茫茫	ボウボウ	々	23.055	1.292	1.414	1.318
98	朦朧	モウロウ		35.062	0.622	2.279	0.763
99	剽軽	ヒョウキン	-	5.379	1.993	3.320	2.744
100	卓袱台	チャブダイ		27.857829	0.990	1.620	1.082

付録 1 百羅漢テスト得点の大学内分布 (近藤と天野、2001 より)

18才から30才までの現役の学部大学生もしくは大学院生(男983名、女424名)
 全体の平均は55.1点、標準偏差は14.8
 但し、得点が15点以下のデータ除外した1410名のデータ
 大学B-E、F-I、J-Mは、大学BとEを除きグループ内で平均値に有意な差がない



付録 2 百羅漢テスト用紙例

No: _____ 名前: _____ (男・女) 実施日: _____ 年 月 日
 年齢: _____ 所属: _____

No	単語	回答欄	No	単語	回答欄	No	単語	回答欄	No	単語	回答欄
1	狸		26	滞日		51	交直		76	鯖	
2	斡旋		27	火蓋		52	唐物		77	算盤	
3	貫祿		28	老衰		53	鯉		78	鰻井	
4	湯煙		29	夜色		54	救医者		79	絆創膏	
5	悠悠		30	鎧		55	細則		80	正骸	
6	湾曲		31	改悪		56	肉饅頭		81	付会	
7	空使		32	髭		57	覇者		82	態態	
8	無恥		33	輔佐		58	塵紙		83	十念	
9	巡業		34	小身		59	水馬		84	汗疹	
10	餌食		35	襟首		60	参画		85	遙遙	
11	民放		36	収賄		61	誰彼		86	青痣	
12	庭草		37	俄然		62	辻褄		87	満配	
13	薄着		38	満帆		63	筏		88	金蔓	
14	出土		39	茶茶		64	鬼瓦		89	笮	
15	曲者		40	磯巾着		65	弓馬		90	目脂	
16	勤続		41	髭面		66	万万		91	誰其	
17	滑稽		42	陰険		67	鳴戸		92	火燵	
18	奴等		43	干菓子		68	蹄		93	黙禱	
19	御日様		44	贅肉		69	梯子		94	徽	
20	亡者		45	急須		70	惚気		95	掠鳥	
21	除数		46	酒粕		71	神州		96	舶載	
22	夜長		47	彼是		72	粗方		97	茫茫	
23	地平		48	天辺		73	竹藪		98	朦朧	
24	執拗		49	腋毛		74	天朝		99	剽軽	
25	蓋		50	奇聞		75	頓知		100	卓袱台	