

岡崎パネル調査の文の長さ

—繰り返しデータの個別表示技法—

Length of Sentences in Panel Surveys of Okazaki

Individual representation of repetitive surveys

(Ver. 1.1)

国立国語研究所

National Institute for Japanese Language and Linguistics

日本語の大規模経年調査に関する総合的研究

Comprehensive Research

Based on Large-Scale, Long-Term Studies of Japanese

井上史雄 ・ 柳村裕

INOUE Fumio, YANAGIMURA Yu

平成 27 年 4 月 29 日

29 Apr 2015

目次

1. はじめに	2
1.1. パネル調査の意義と成果	2
2. パネル調査の位置づけ	2
2.1. 住民構成の変化の可能性	2
2.2. 追跡調査パネルサンプルのまとめ方	3
3. パネルサンプル 結果の概要	4
3.1. パネル調査表示の文の長さ 全話者	4
3.2. パネル調査表示の文の長さ 第1/2次話者	7
3.3. パネル調査表示の文の長さ 第2/3次話者	8
4. 長さの対比とパネルデータの再検討	9
4.1. 文字化の忠実度の影響	9
4.2. 全データの生年実年代グラフ	9
4.3. 住民構成の変化の影響	11
4.4. 項目別回答文の長さ パネルデータ	11
5. 長さ一覧資料（各項目）	13
101 道教え	14
102 振込用紙	15
103 荷物預け	16
104 傘忘れ	17
105 先生	18
106 新聞代	19
107 議事堂	20
108 医者	21
109 席譲られ	22
110 おつり	23
111 傘貸し	24
6. 結論 現代の言語変化	25
参考文献	25

1. はじめに

本編は、「日本語の大規模経年調査に関する総合的研究」の逐次報告であり、国立国語研究所がこれまで半世紀以上にわたって継続した岡崎敬語調査に関する資料集である。今回はパネル調査（同一人物追跡調査データ）の表示法について報告する。

1.1. パネル調査の意義と成果

パネル調査、つまり同一人物追跡調査データは、変化過程を知るには有用である。ただしパネルサンプルの研究技法は確立したとは言い難い。信頼できる結論を得るためには大勢のデータが必要ではあるが、追跡調査が可能だった話者には年齢や性別の偏りがあり、平均値だけをもって論じることは危険でありうる。一方個々のケースを個別に論じるのは手間がかかるし、一般性を見失う可能性がある。この稿では、一方の軸に全話者を生年月順に配置し、もう一つの軸に言語事象の数値をプロットして、データ全体を検討する技法を紹介する。

初期の成果の速報のために、岡崎調査（国立国語研究所 1958, 1983, 2010）の文長（反応文、回答の長さ）を分析対象にした。これにより、一人ひとりのデータでも、生涯変化（加齢変化）として反応文が長くなることが分かった。また場面別に見ることにより、どの場面の増え方が大きいか分かった。さらに個々の表現別に見ることにより、どの言い方がどんな人で増えたかも分かった。

岡崎の反応文の長さは、話者全体のデータできれいな成人後採用を示した。「年寄りの長話」と呼べる。これは「敬語の成人後採用」とも関連する。教科書での敬語の扱い方や、社会人向けの敬語教育にも影響するところが大きい。また学校敬語の特殊性の位置付けにも関係する。

2. パネル調査の位置づけ

2.1. 住民構成の変化の可能性

岡崎で得られた成果は、住民全体についてだった。岡崎 3 回の調査結果の違いについては、住民構成の変化、「よそもの」の流入の影響によって生じたという可能性がある。山形県鶴岡市、愛知県岡崎市のように数十年をへだてて大規模サンプリング調査を行った場合の実時間の違いは、移転などによる住民構成の変化を反映する可能性もあり、ニュータウンにおける方言形成と似た状況も想定できる（朝日 2008）。ことに岡崎市では、隣接の豊田市という工業都市の住宅地として、岡崎市に流入する住民が増えた。鶴岡市のように首都圏への人口流出（と後背地農山漁村からの流入）が主で、減少傾向にある地域とは性格が違う。

純粹の言語変化を扱うには、同一人物追跡によるパネル調査の結果を重んじるべきである。また個人内の加齢による変化を考慮すべきである。同一個人が数十年経って年取ったときに同じような（長く話す、答えるという）傾向を見せるだろうか。岡崎にはこの貴重な追跡データがある。

2.2. 追跡調査パネルサンプルのまとめ方

岡崎敬語調査の調査デザインと規模は図1に要約できる。ランダムサンプルをその都度住民票から抽出するとともに、パネルサンプルとして、以前の話者を探し出した。

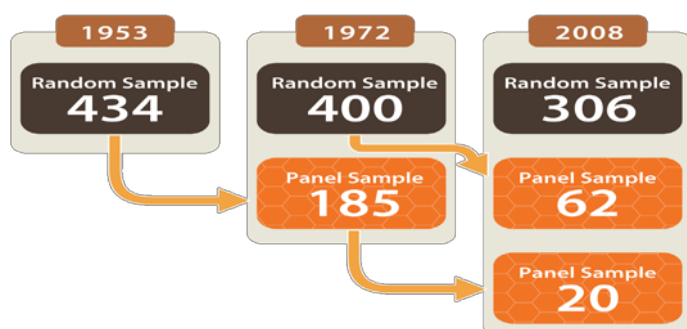


図1 岡崎パネルデータの構成

表1 岡崎パネルデータ 調査当時の年齢

生年	①1953年調査	②1972年調査	③2008年調査
1880	①73歳	②92歳	
1890	①63歳	②82歳	
1900	①53歳	②72歳	
1910	①43歳	②62歳	③98歳
1920	①33歳	②52歳	③88歳
1930	①23歳	②42歳	③78歳
1940	①13歳	②32歳	③68歳
1950		②22歳	③58歳
1960		②12歳	③48歳
1970			③38歳
1980			③28歳
1990			③18歳

図1の読み取りの参考として表1に10年齢層ごとに、3回の調査のときの年齢を示す。1930年、1940年ころに生まれた人は、第1次から第3次までの3回の調査に協力できる年齢だった。図1によれば実際には20人だった。

パネル調査の問題点は、あとになるほど人数が少なくなることである。統計的有意差などを計算しようとする信頼度が落ちて、結果が不分明になる。また年齢などによるグラ

フを作っても、各セルの人数が少なく、一定のパターンが現れにくい。この弱点を逃れる（またはカバーする）ために個々のケーススタディーを行うことがあるが、これはまた一般化が困難になるという欠陥がある。

以上のジレンマを逃れる1技法として、エクセル散布図の利用が考えられる。個別例を図表の形で配置するもので、基本的発想・根本精神としては、言語地図やグロットグラム（地理×年齢図）と同じである。少数例の平均値でなく、個人表示をするのは、グロットグラムの発展である。岡崎市を1地点とみなし、年齢の軸を細分して、生年月日により1次元に配列したことになる。残った1次元に言語現象の使用率を表示した。

この場合、個々の例を絶対的な不動のものとして見るのではなく、全体のパターンをとらえることが必要である。言語地図やグロットグラムの個人表示では、全体のパターンを読み取る必要があり、名人芸がときに要求されるが、幸いに今回は連続的数値を扱う。数表でなくグラフで視覚化したので、表面的な解釈に個人差は少ない。ただし究極的な論理の読み取りの深さには経験による違いや個人差がありうる。

3. パネルサンプル 結果の概要

3.1. パネル調査表示の文の長さ 全話者

これまで全体パターンの類型化を試みて、「S字、IIの字、=（イコール）」（旧名「S字、二の字、川の字変化」）という三つの変化パターンを提示した。以下では岡崎敬語調査の中で思いがけなく興味深いパターン「イコール変化」を示した現象の典型、回答文の長さについて扱う。3回の調査全体を通して長くなる傾向が見られ、にも関わらず年齢差を見ると若い世代では短くなる傾向がある。左上がりの線が繰り返される。実時間の繰り返し（経年）調査としては長くなり、年齢差と言う見かけの時間では短くなるという、矛盾した方向が示された。同一個人の生涯を通して追跡したらどうなるだろうか。

図2に散布図の技法により提示する。縦軸が話者の生年を示す。月まで考慮することにより、全員を判別できた。横軸は文の長さをモーラの数で示す。この表示法は連続的数値について適用可能である。今後丁寧さの段階付け、「テモラウ、テイタダク」使用率、ヘッジ使用率、文体（テゴザイマス、ため口使用率）などにも拡大する予定である。一方非連続の名目変数にあたる1/0データについては、「●△ー」などの記号を使って有無を表示する技法で、1次から3次にかけてのパネル話者の変化を表示できる。グロットグラムや社会言語学の一覧表で使われていた技法の応用である。

以前の報告は文字数を用いたために、拗音の部分で1単位長くなり、「チョット」や「ミョーダイバシ」などで問題が生じたが、全体の文字数から見ると、問題は少ないとみた。今回の集計ではモーラで数えたので、拗音の問題がなくなった。以前のグラフや集計と比べても差が見られないので、以前の考察に変更を加える必要はない。

第1次から第3次までの調査次は数字のラベルで示した。ただし第1次の数字は省いた。これにより、線を矢印のように読み取ることができる。大部分の線は右向きである。つまり後の調査で確かに文が長くなる。一部短くなるケースは、線を点線に変えた。1900年以前の生まれの人（第1次調査ですでに高齢だった人）に多い。ある程度年を取ると、文がさらに長くなる傾向は収まるものらしい。

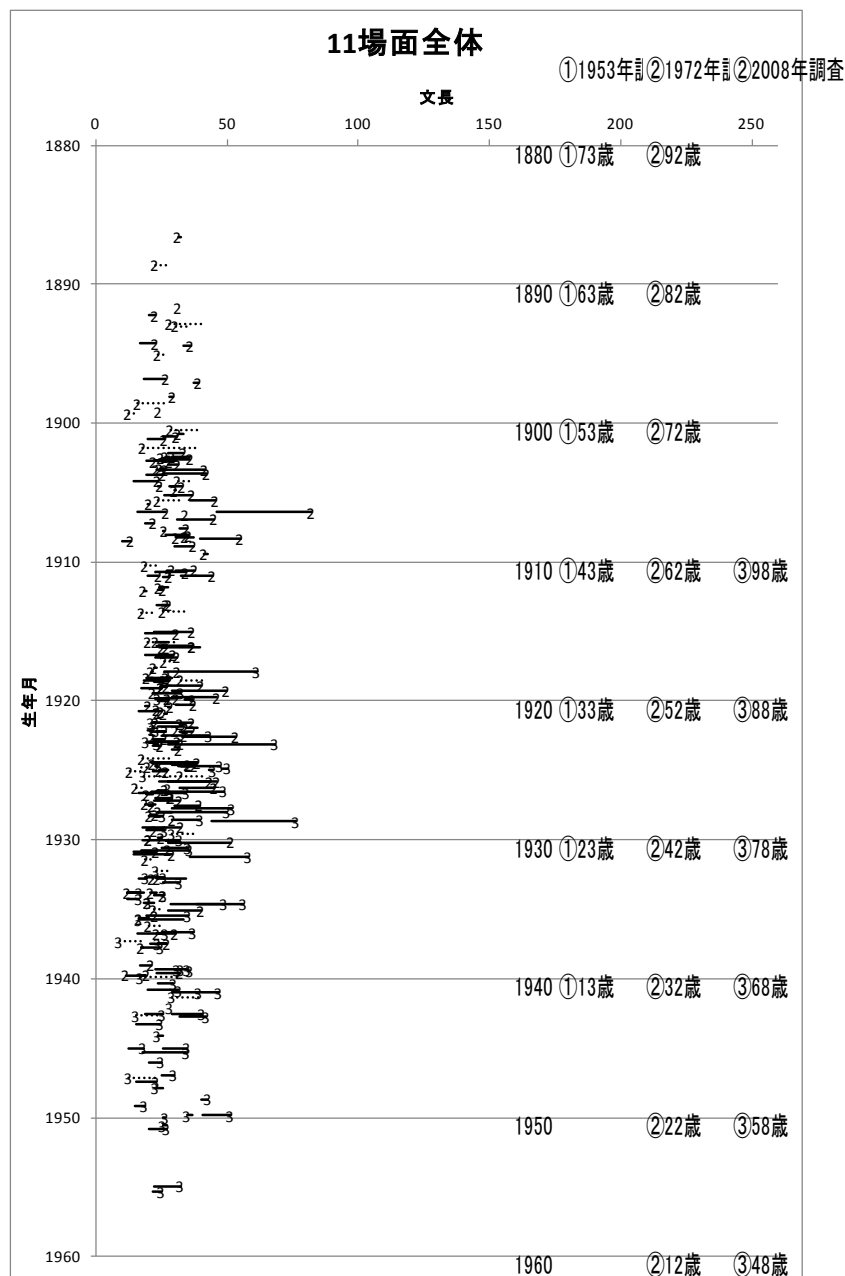


図2 岡崎パネル全データの文の長さ

図3に図2の主要部分を拡大して示す。同一個人を19年後、36年後に追跡して調べたところ、全体としては、回答文が長くなる人が圧倒的に多いことが示された。

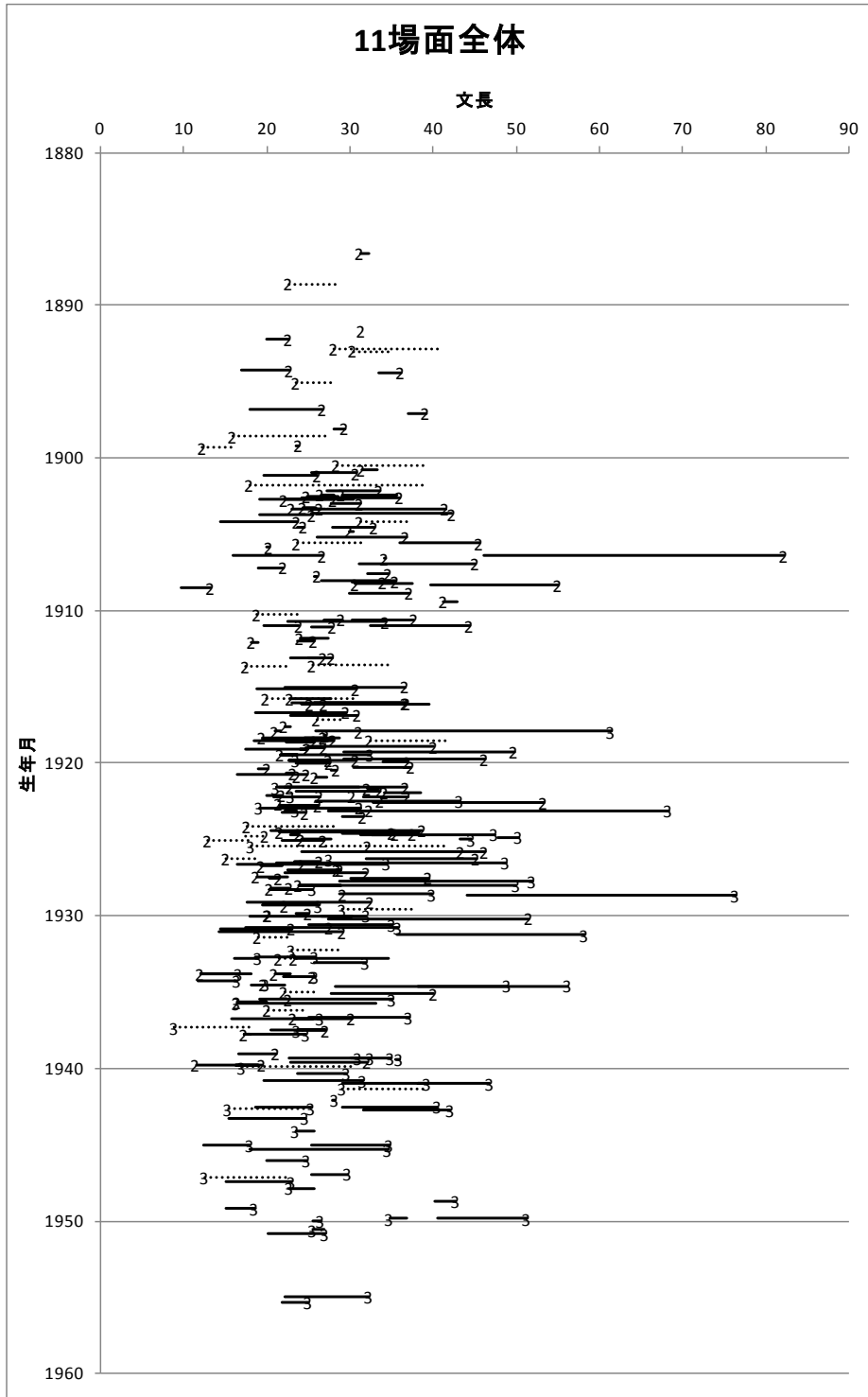


図3 岡崎パネル全データの文の長さ (拡大)

3.2. パネル調査表示の文の長さ 第1/2次話者

図2, 3のグラフのデータは3回の調査のパネルデータ全体を含む。細分したらもっと傾向が読み取りやすくなると期待される。図4a, bに話者を2グループに細分して、岡崎調査の文の長さを4枚に分けて示す。対比して見やすいように90度傾けた。個々の話者を表示したので、個別のケースも把握できるが、全体的傾向を要約して示す必要もある。ここでは年齢層に分けて平均値を出す手法は避けた。パネル調査では人数の少ないセルが生じがちで、個人の偏った回答の影響が大きく働きうるからである。代わりにグラフ中に近似直線と近似曲線を入れて、傾向を把握する。

図4aの2枚は1次調査全員のうち2次のパネル調査の対象になった話者だけを取り上げたもの(左)と2次調査のパネル話者のデータ(右)である。近似直線と近似2次曲線を入れた。同一人物かどうかは左右の生年で判別できるが、今回は左右の幅を狭めたので、同定が困難である。(図2, 3で分かることである。)左の第1次調査と右の第2次調査を比べると、パターンが違い、第2次で長い回答文が増えた。また近似直線から分かるように、第2次調査全体で文が長くなった。中でも長いのは中年層で、高齢者と若い人は短い。

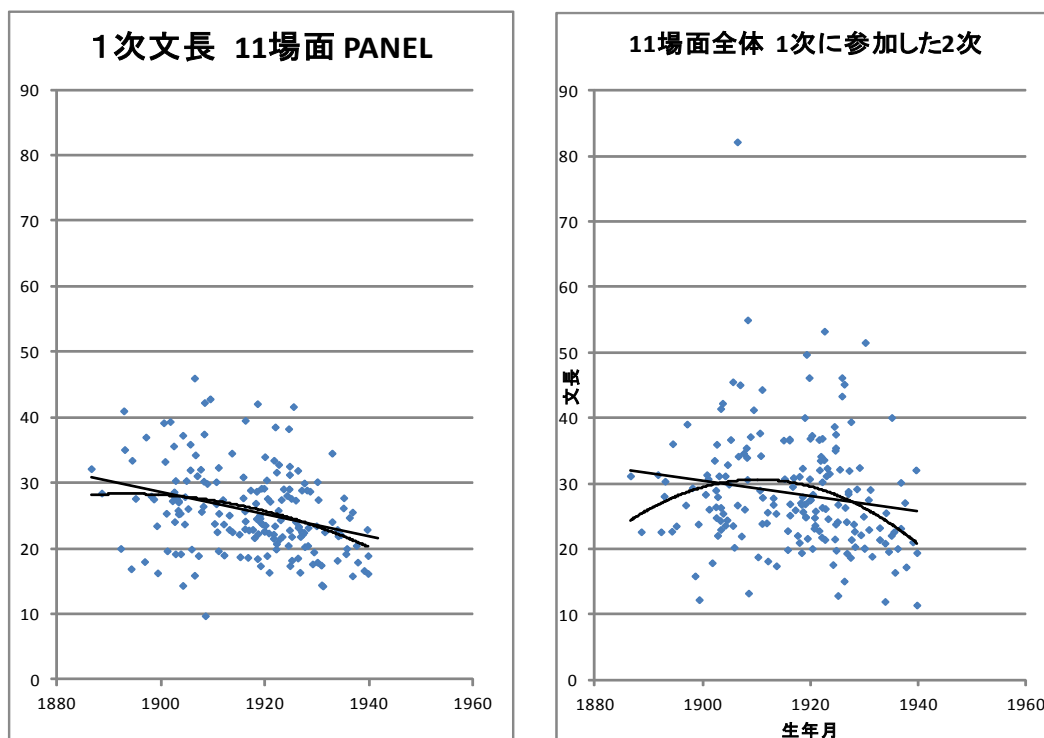


図4a 岡崎調査の文の長さ 1次パネル・2次パネルの比較

3.3. パネル調査表示の文の長さ 第2/3次話者

図4bの2枚は2次調査全員のうち3次のパネル調査の対象になった話者だけを取り上げたもの（左）と3次調査のパネル話者のデータ（右）である。1次から3次まで参加している20人は図4bのグラフと重複して図化されている。左と右のグラフを比べると、ここでも後の調査で文が長くなる傾向が見られる。近似直線の位置から、第2次調査全体で長くなったと読み取れる。近似曲線は上の2枚と違って凹型で、老年層が長く、また若年層もやや長い。第2次と比べて全体として文が長くなっている。

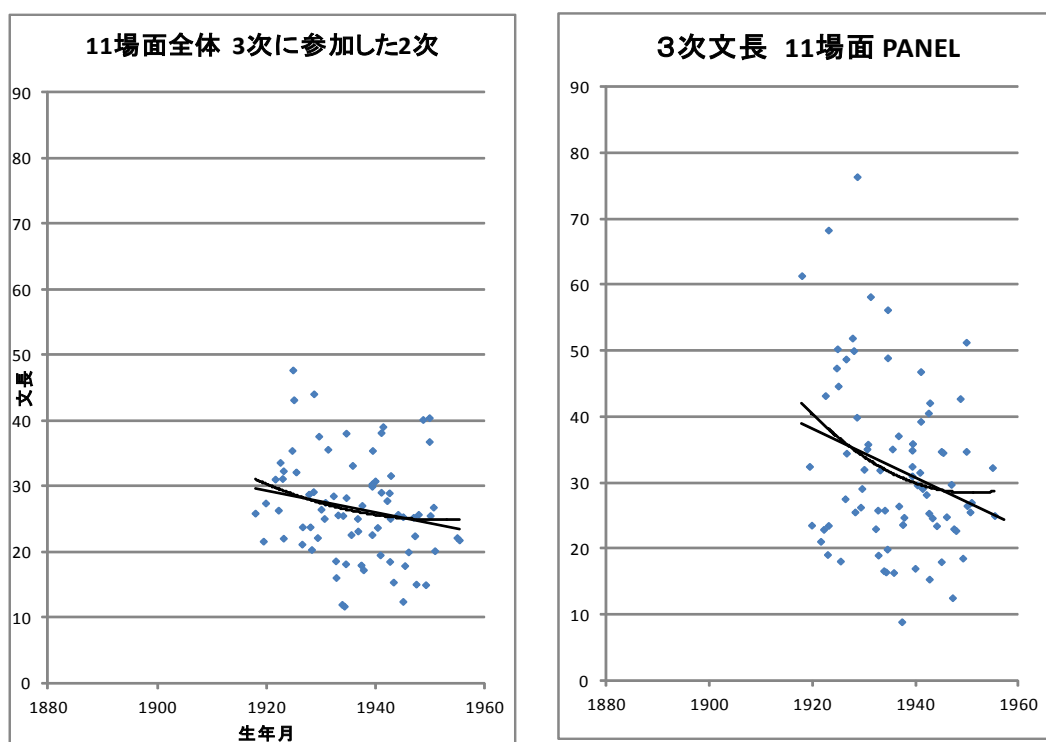


図4b 岡崎調査の文の長さ 2次パネル・3次パネルの比較

以上で、文が長くなるのは、岡崎の住民構成が変わった（よそものが増えたために、例えば道案内のときに詳しく説明するようになった）という理由ではないと推論できる。同じ人が19年後、36年後に同じ中身を伝えるのに長く話すようになったことが示された。これこそ同一人物を追跡して行うパネル調査の価値である。実時間の差が分かるし、個人の経年変化、加齢変化が実証できた。しかもかなりの（100人前後の）人数の全体的傾向が把握できた。

以上図4a, bで、「イコール変化」の右下がり（左上がり）の線に着目して、岡崎敬語データから得られた4個のグラフを分析した。相互に補強しあう形で、敬語関連現象の成人後採用という一般傾向を示した。

4. 長さの対比とパネルデータの再検討

4.1. 文字化の忠実度の影響

3回にわたって回答文が長くなった理由について、実際に文が長くなったのではなく、文字化の忠実度が影響したという反論がありうる。1953年の第1次調査では録音機を使用せず、その場の短期記憶を手書きで書きとった。ベテラン調査員 Professional と愛知学芸大学学生 College で技術と考え方の差があったと考えられる。それに反して、第3次の文字起こしでは、言い淀み、ヘッジに関わる部分も忠実に文字化しようという方針を取ったそうである。その影響はありうるが、各調査次の全員、全年齢層に同様に働くはずである。確認のために、年齢差を検証する。

4.2. 全データの生年実年代グラフ

図5に、これまでの3回の調査データを6本で表現したグラフを示す。結論を先取りしうると、その読み取りによれば、たとえ文字起こしの技術と方針に差があったとしても、「成人後採用」の傾向を崩すものでない。

図5のグラフの横軸は話者の生年、縦軸は回答文の文字数を示す。6本の線のうち、点線は第1次調査、破線は第2次調査、実線は第3次調査を示す。太い線は各調査次で重視すべきデータで、第1次のプロパーグループ（または Professional 調査員）、第2次・第3次のランダムサンプルである。（第1、2次調査の報告書で扱われている第1次調査のデータは、実はプロパーグループのみである。今後これを「第1次ランダムサンプル」と呼び、コントロールグループを「第1次コントロールグループ」と呼ぶことにする。）細い線は参考にするデータで、第1次のコントロールグループ（または College 調査員）、第2次・第3次のパネル（追跡調査）サンプルである。それぞれのセルの人数は十分である。ただし第3次パネルの90代は一人だけだったので、80代に含めた。このため第3次パネルの80代は、各年代の話者数の比率に応じて、横軸上でわずかに左にずらして表示した（本来、80代は1924年、90代は1914年にプロットされる。80代が21名、90代が1名なので、 $(1924 \times 21 + 1914 \times 1) \div 22 = 1923.5$ 年にプロットした）。

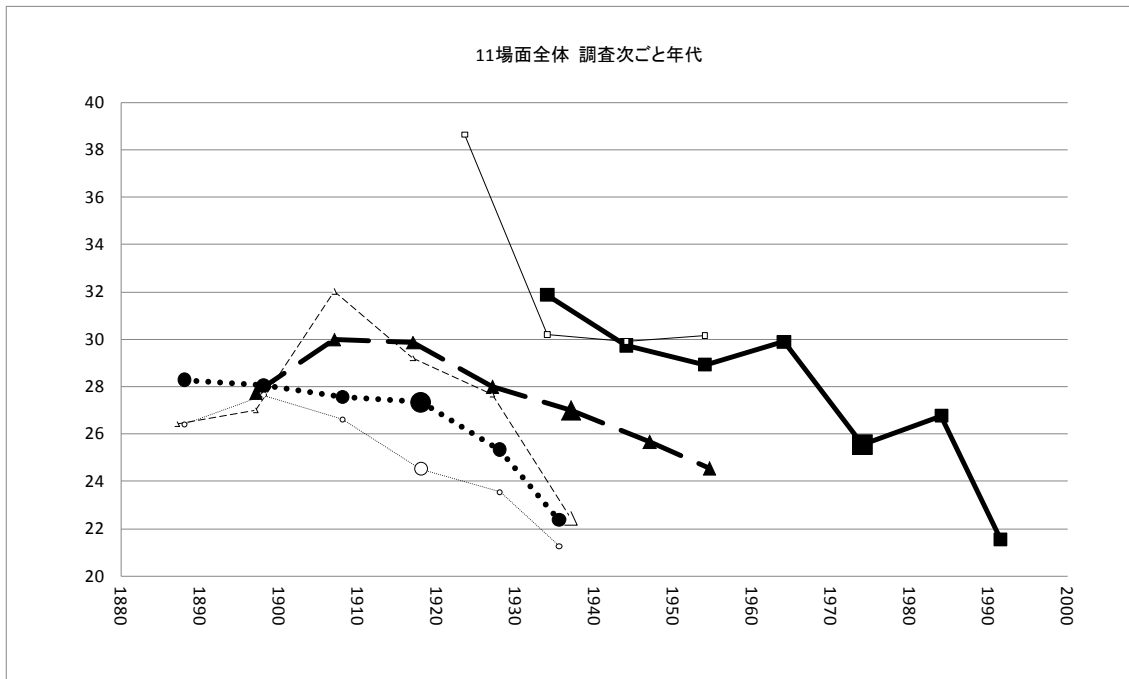


図5 岡崎敬語調査 文長平均使用数 11 場面全体調査次ごと年代

3本の太い線はこれまでに生年実年代グラフで示したものと同一である。年齢差を見ると、3回とも左上がり、成人後採用パターンを示す。細い線は、太い線と少しずれることがある。そのうち点線の第1次 College 調査員は Professional 調査員より短い。書き取り能力・忠実度の違いを示す可能性があるが、一方で、話者が若い地元の学生に気を許して、くだけた感じの回答を口にした可能性もある。細い線のうち破線の第2次調査のパネルサンプルは、その前のランダムサンプルの一部である。(なお図4 a, b で表示された話者は、パネルサンプルだけである。すなわち第1次調査のパネルサンプルとは第1次調査の Professional と College の全ランダムサンプルのうち、第2次調査でパネルサンプルになった人だけをとりあげてプロットした。第2次と第3次調査の関係も並行的である。)

パネルサンプルは追跡に成功した人々で、直前の転入者ではない。この人たちは最高年齢層を除くと、以前の調査次よりもすべて長く回答している。また第2次と第3次のパネルサンプルとランダムサンプルを比べると、年齢層によりずれが見られるが、そう大きいものではない(38年後の調査なので、重なる年齢層は少ない)。各調査次において、(第1次・第2次調査の高年齢層を除くと)左上がりの線を描くことは、全体に適用できる。文字化の方針の違い、変化によるとすれば、第1次から第3次にかけて全年齢層にわたって数値の大きな跳ね上がりが観察されるだろうが、それほどではない。

従って、3回の調査で文字化の方針の違いがあったとしても、年齢層の違いに基づく成人後採用の働きは、否定できない。

4.3. 住民構成の変化の影響

以下では、回答文の長さを 11 項目に分けて観察する。技法としては上記と同様である。3 回にわたって回答文が長くなった理由について、住民によそのもの（転入者）が増えたために 101 道教え項目で長くなったのが主要因だという解釈があった。住民にとって場所が分かりにくくなったために、長く説明するようになったという論理である。確認のために、その他の項目でも長くなったかどうかを確認する必要がある。

全 11 項目の比較のために、図 6 に、ランダムサンプル全員の回答文の長さを示す。赤字は特に長くなった 3 項目である。図 7 を要約して示すことになる。図 7 のグラフと照合すると、ランダムサンプルもパネルサンプルも同様に長さを長くしていると読み取れる。

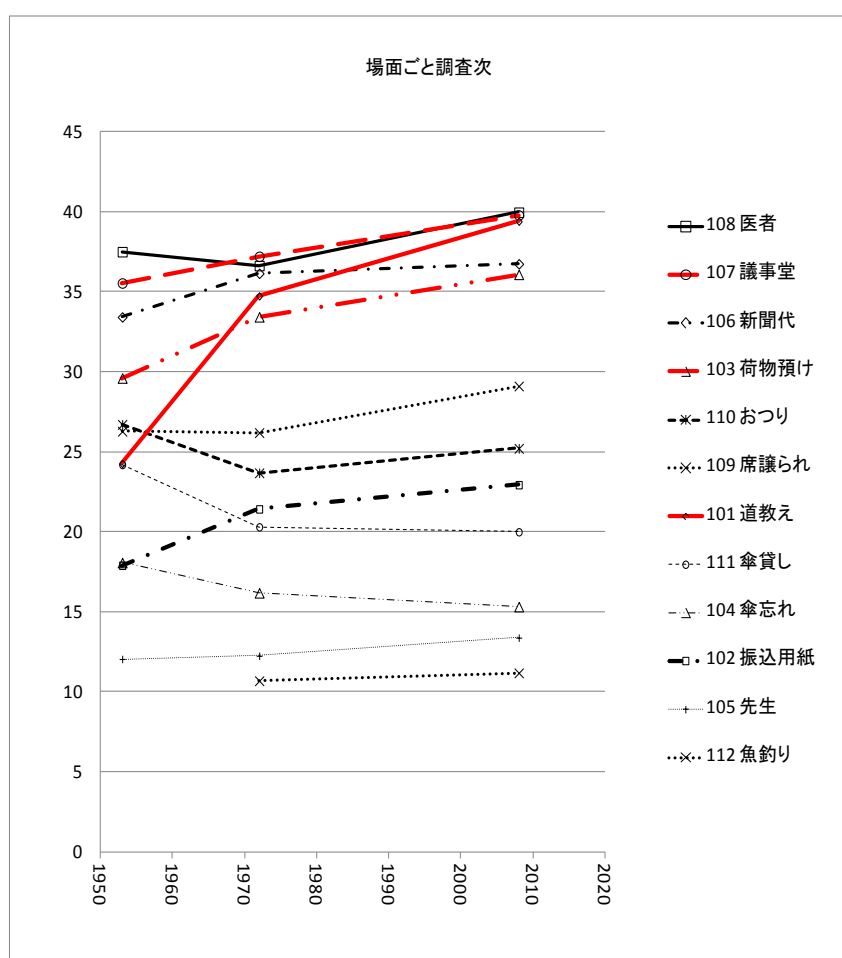


図 6 岡崎調査の文の長さ ランダムデータの比較

4.4. 項目別回答文の長さ パネルデータ

パネルデータの全 11 項目を 3~4 枚並べて相互に比較する。

確かに 101 道教え項目で長くなった。また議事堂、傘貸しなどでも長くなった。しかし全項目で長くなる現象が観察された。よそのもの影響などでなく、ほぼ全項目にかかわる

言語的要因が働いていると考えられる。

101 道教え 102 振込用紙 103 荷物預け

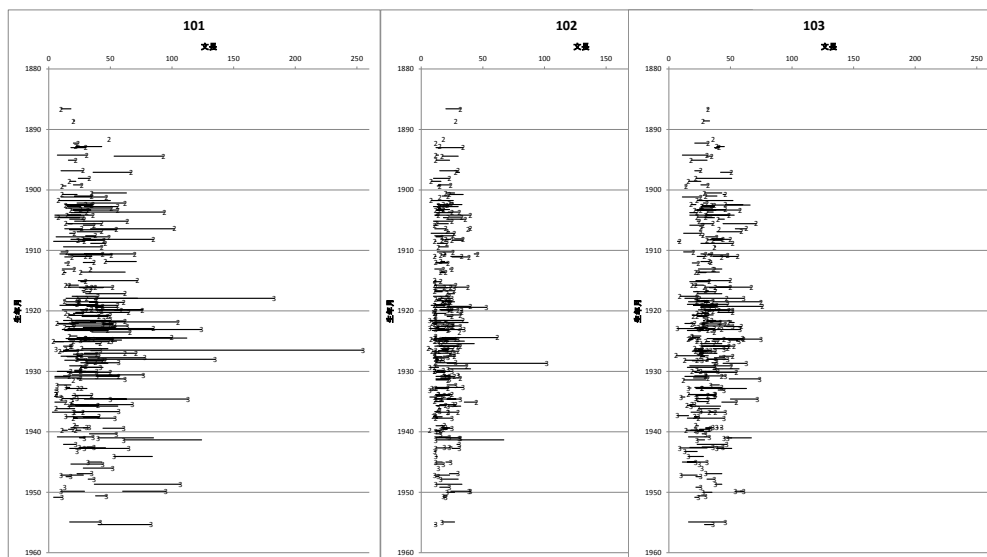


図 7 a 回答文の長さ 11 項目全体 (パネルサンプル 全員)

104 傘忘れ 105 先生 106 新聞代 107 議事堂

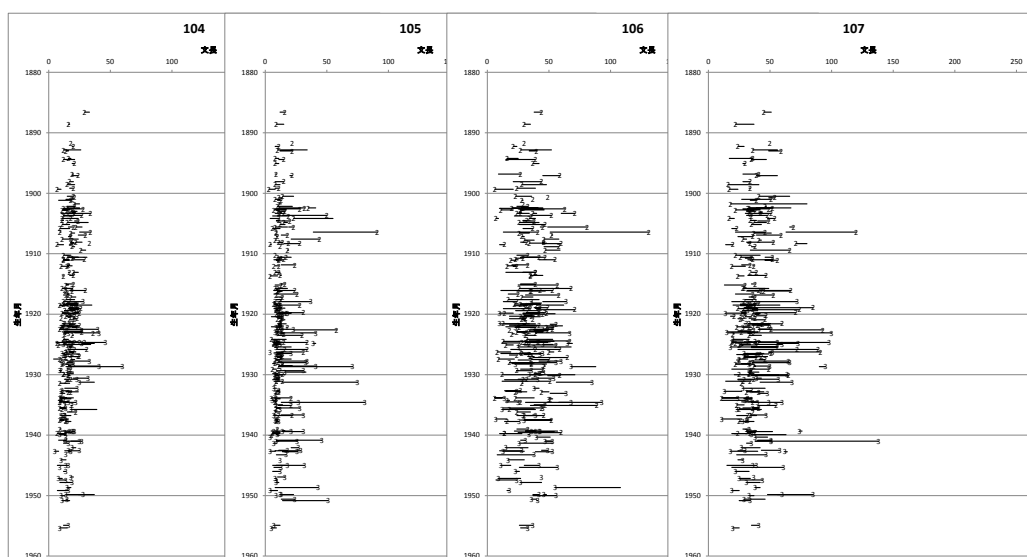


図 7 b 回答文の長さ 11 項目全体 (パネルサンプル 全員)

108 医者 109 席譲られ 110 おつり 111 傘貸し

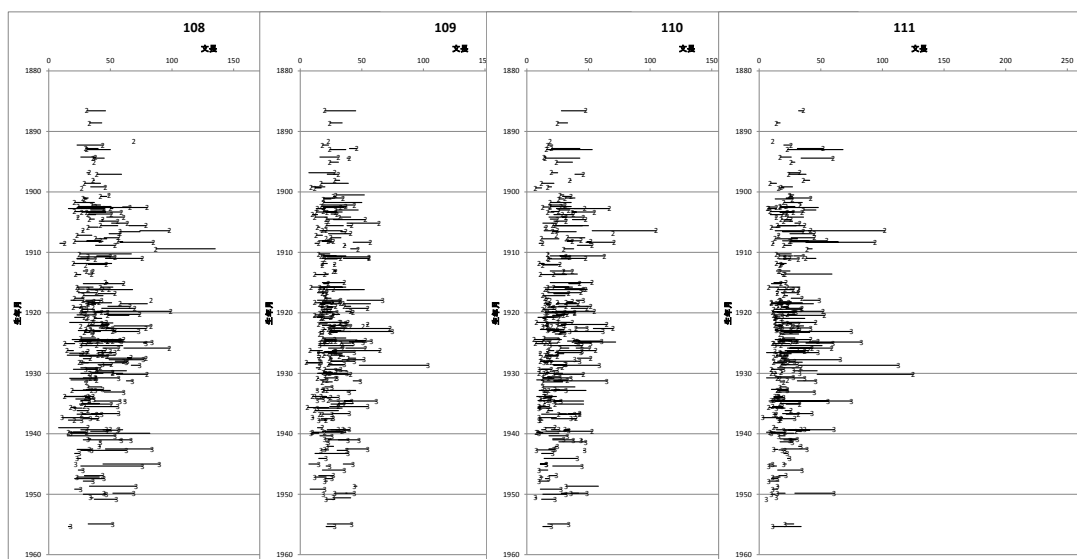


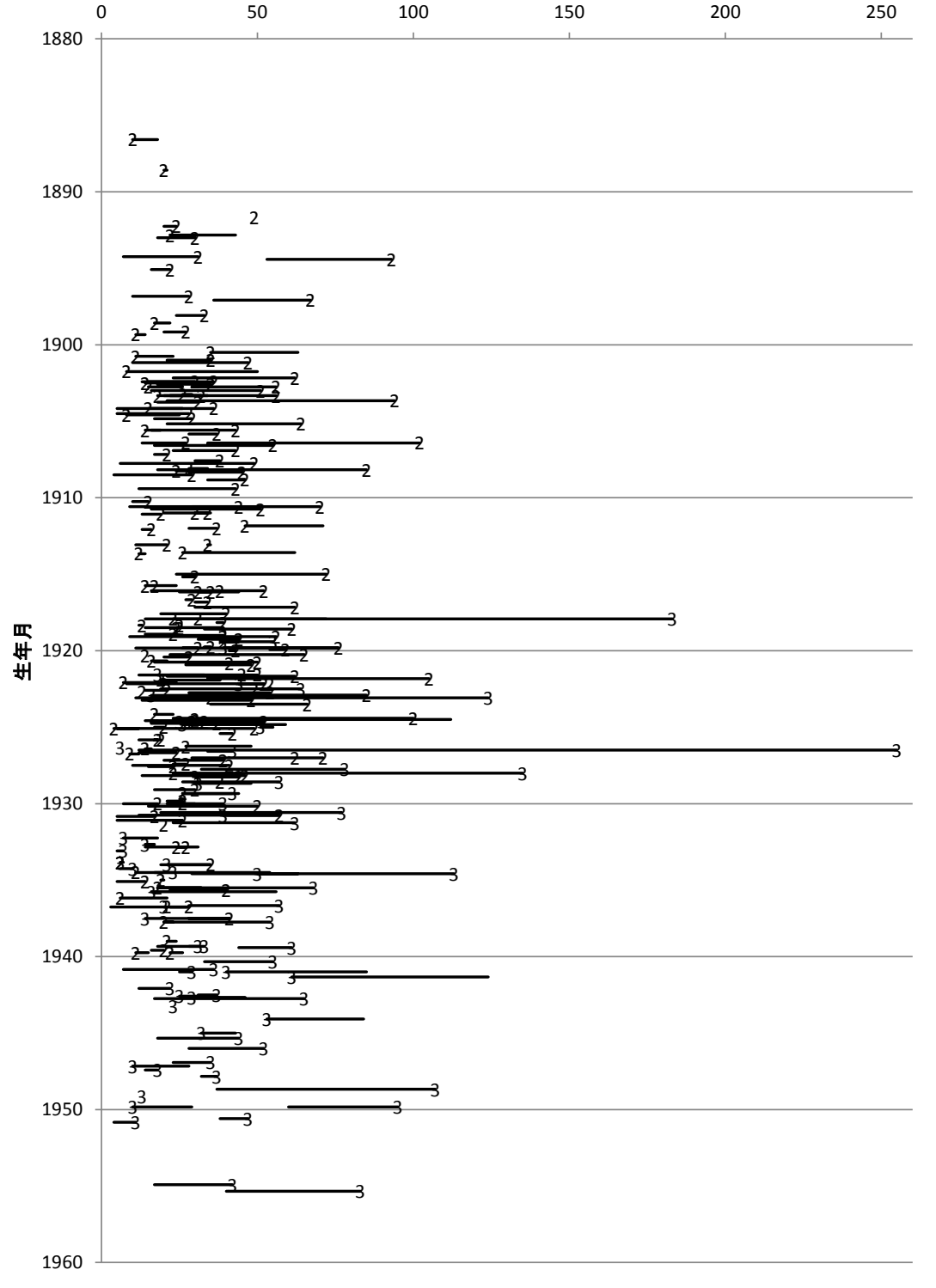
図7c 回答文の長さ 11項目全体（パネルサンプル 全員）

5. 長さ一覧資料（各項目）

以下では11項目それぞれの散布図を提示する。矢印に代わる2, 3の数字で、長くなったかどうかの判別をするためには、大きな図版が必要である。なお図2, 図3と違って、短くなった個人を点線に変えるという手作業はほどこしていない。

101

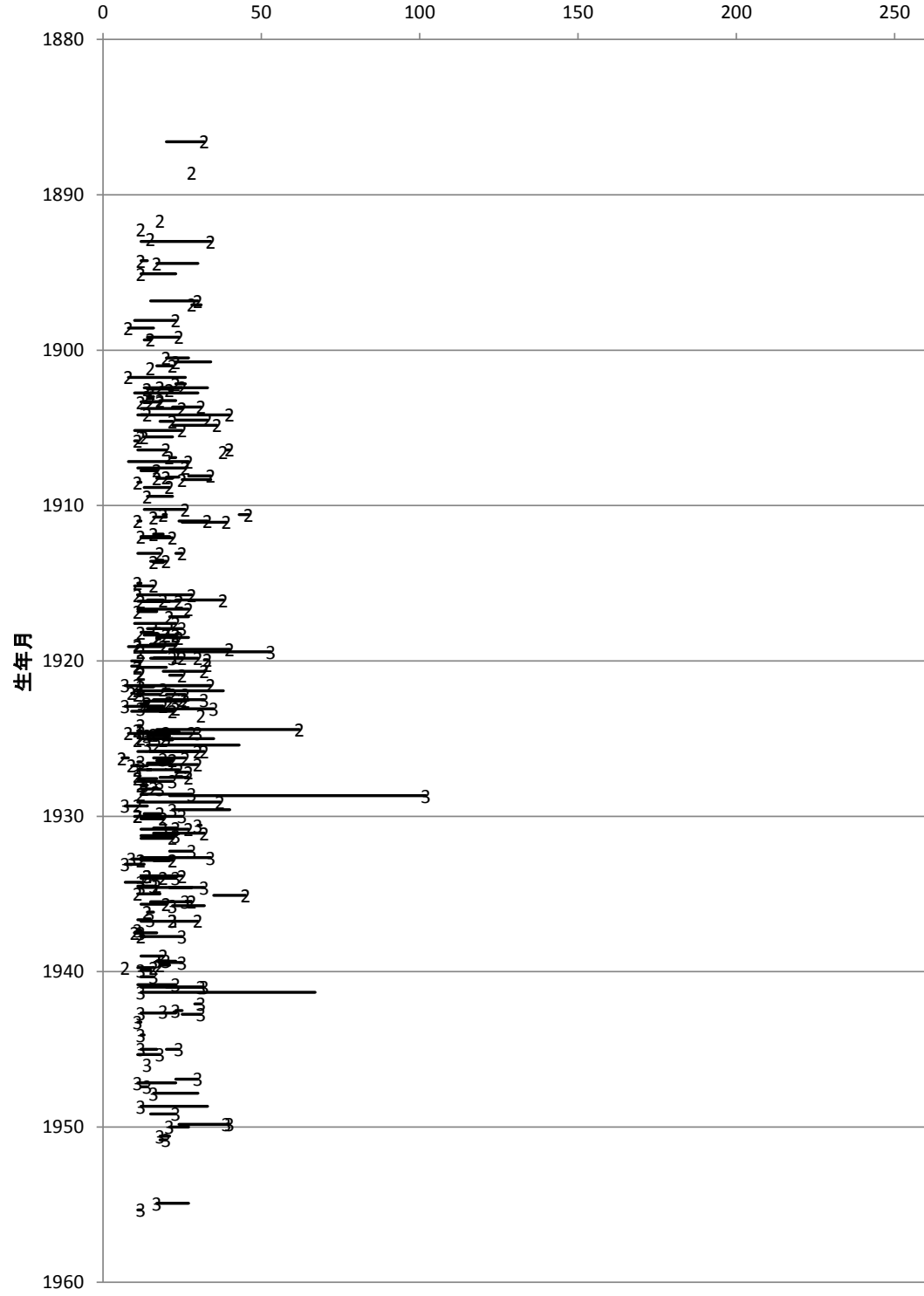
文長



101 道教え

102

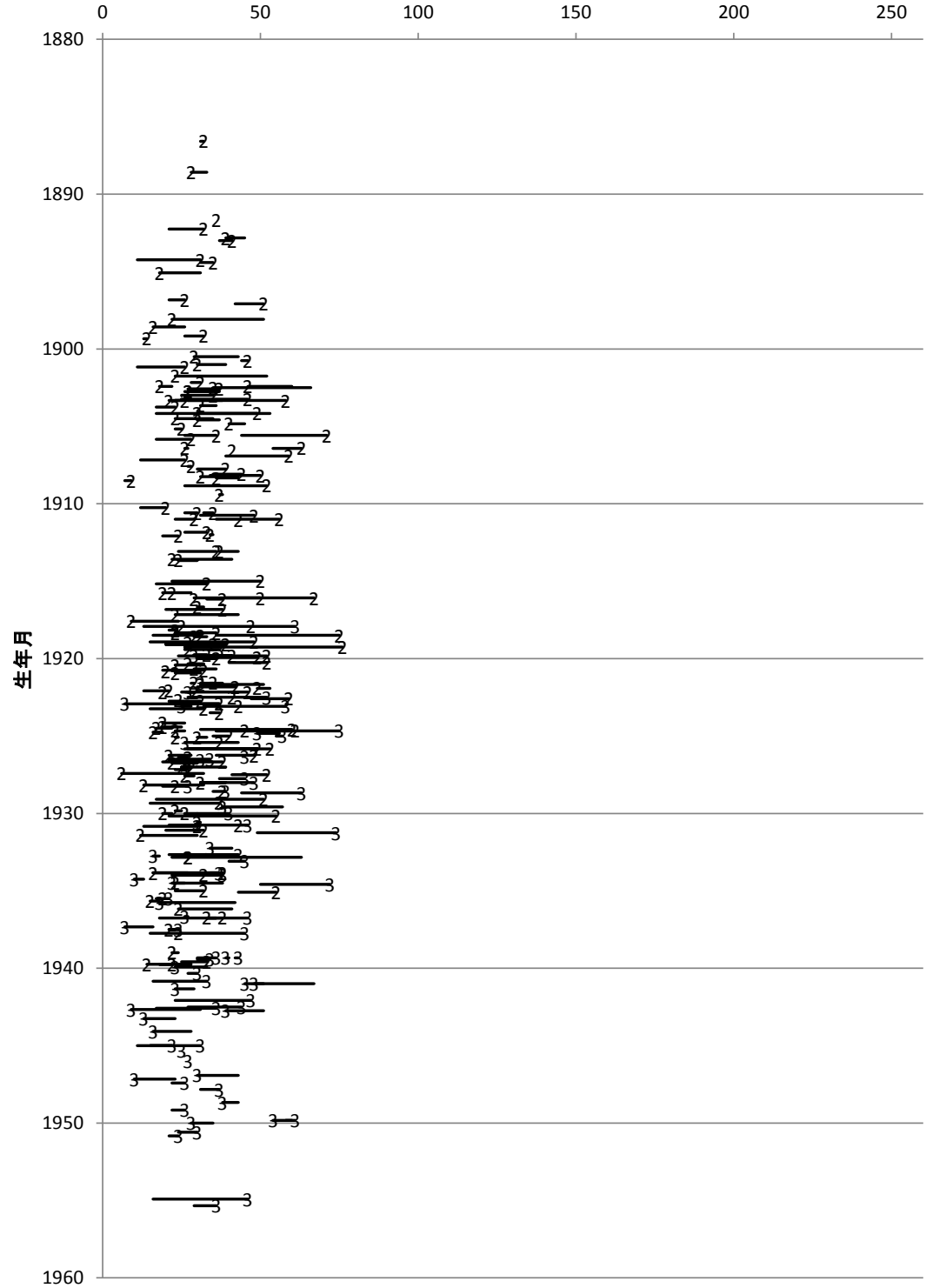
文長



102 振込用紙

103

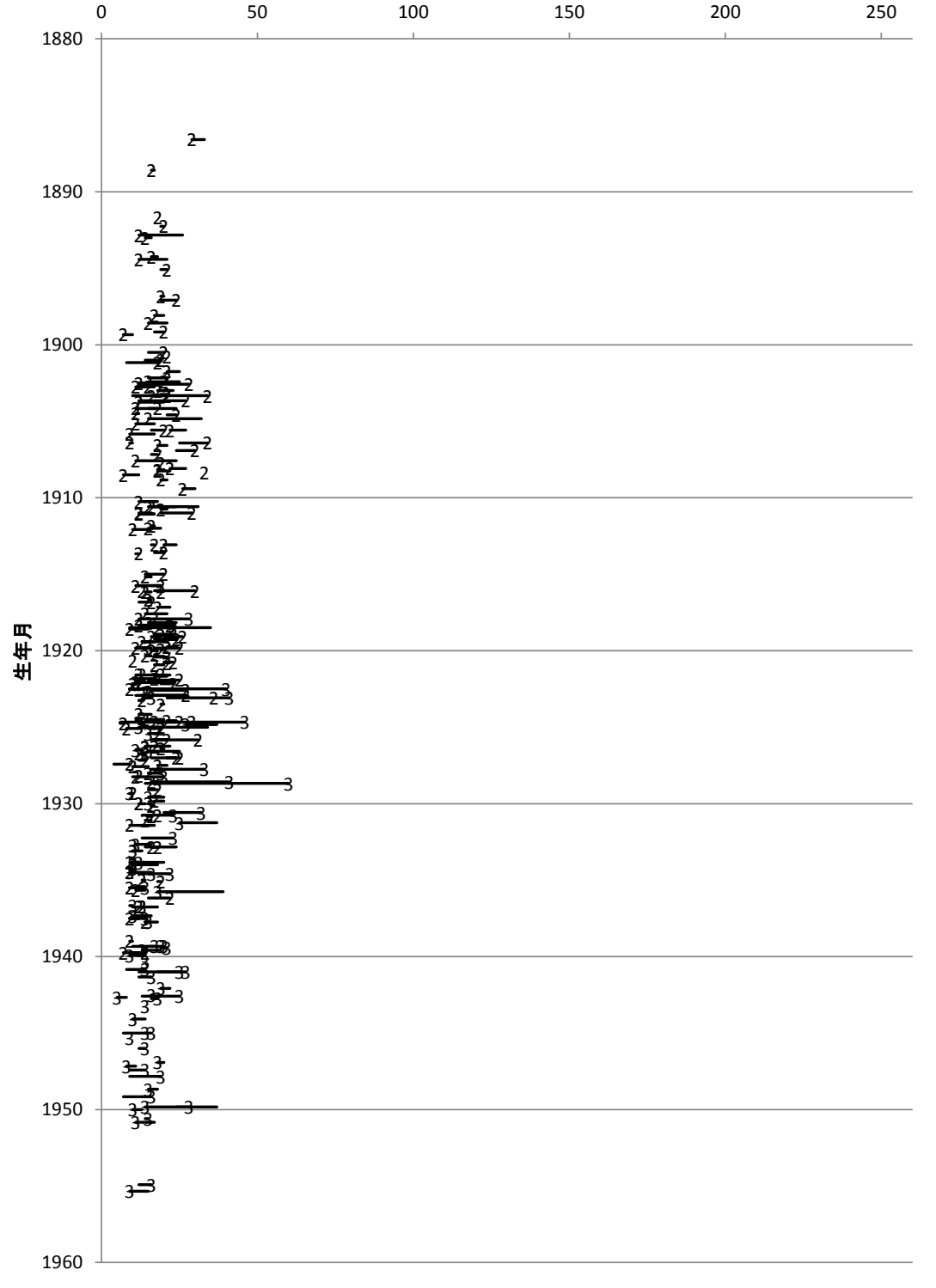
文長



103 荷物預け

104

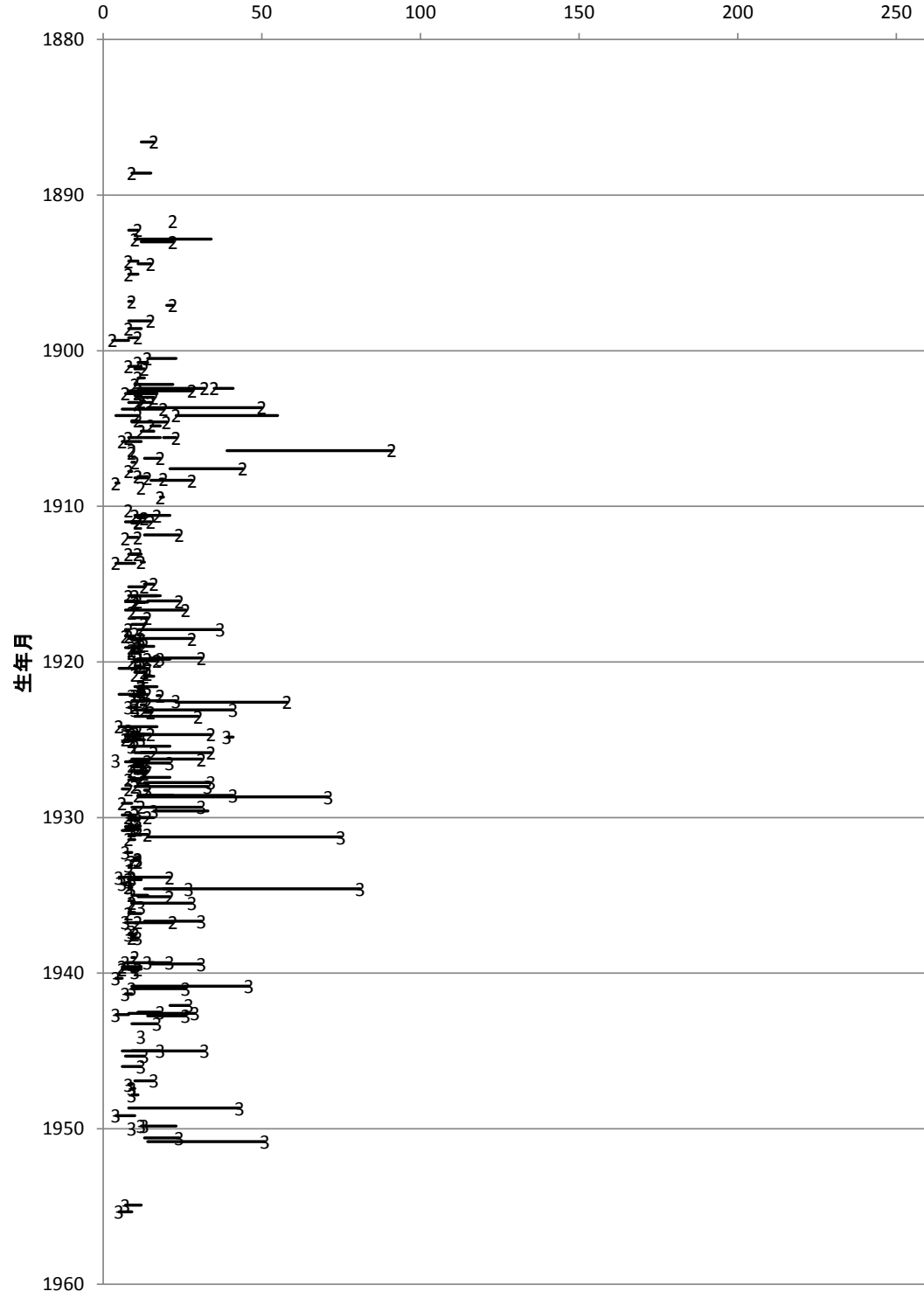
文長



104 傘忘れ

105

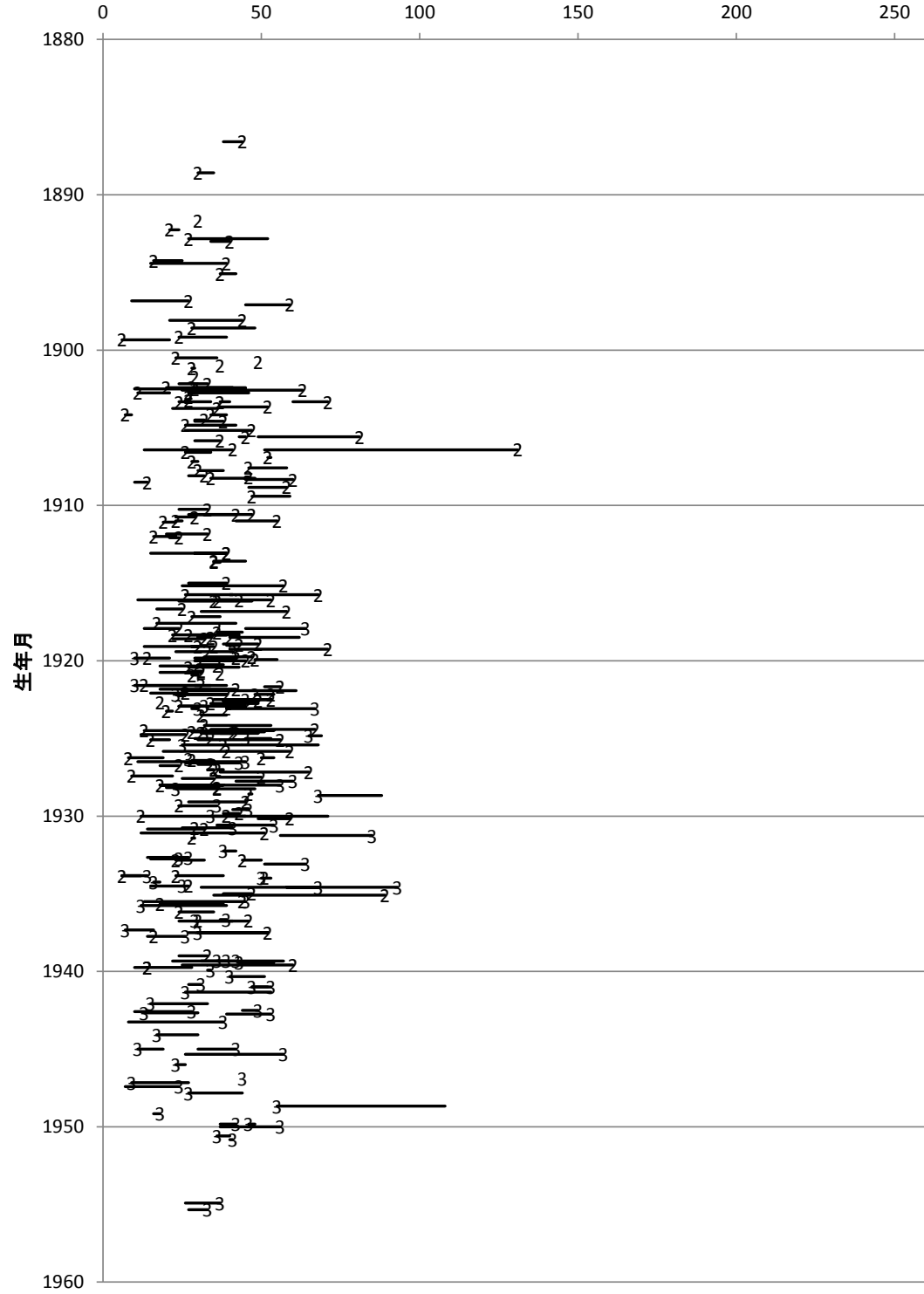
文長



105 先生

106

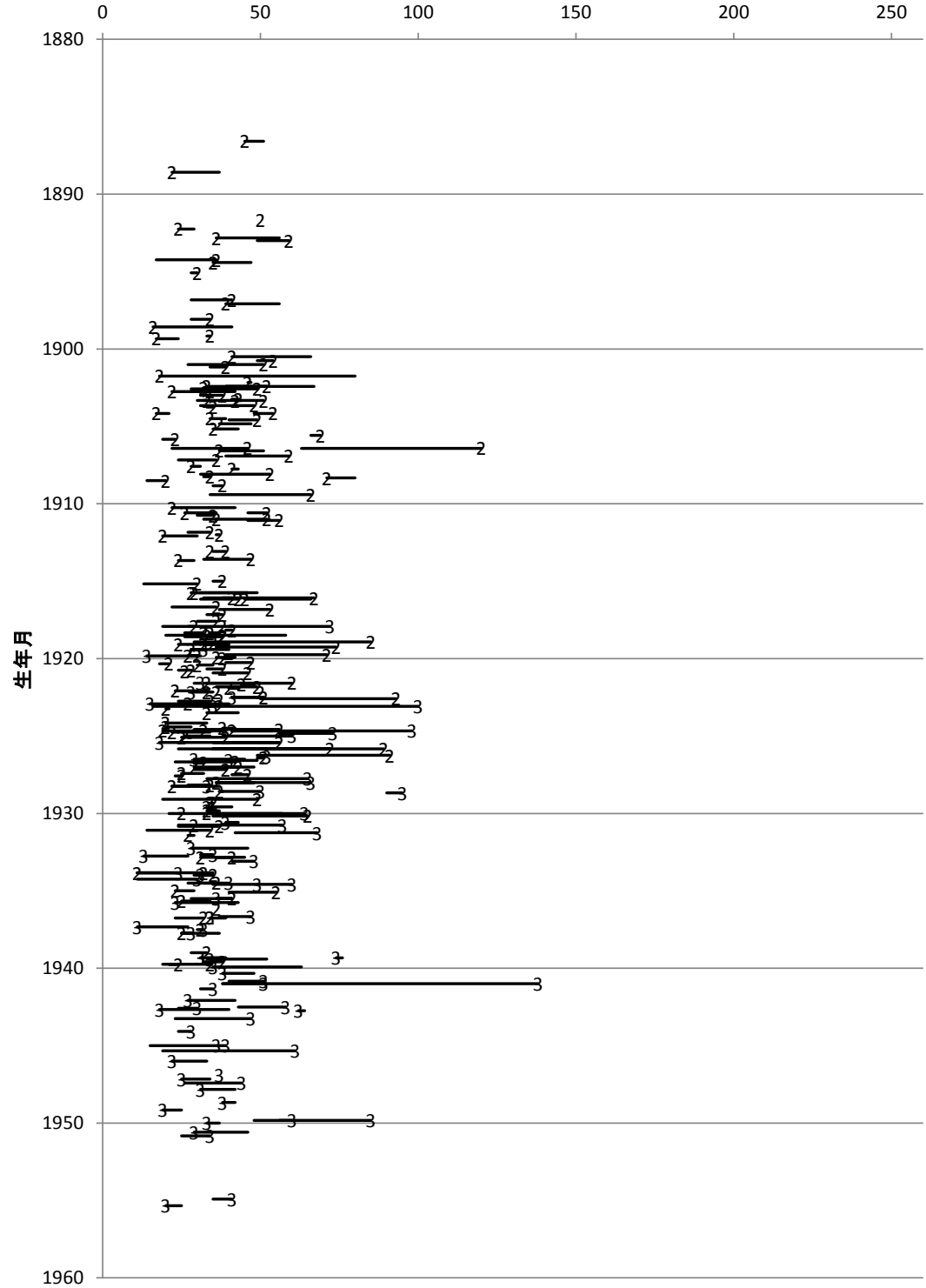
文長



106 新聞代

107

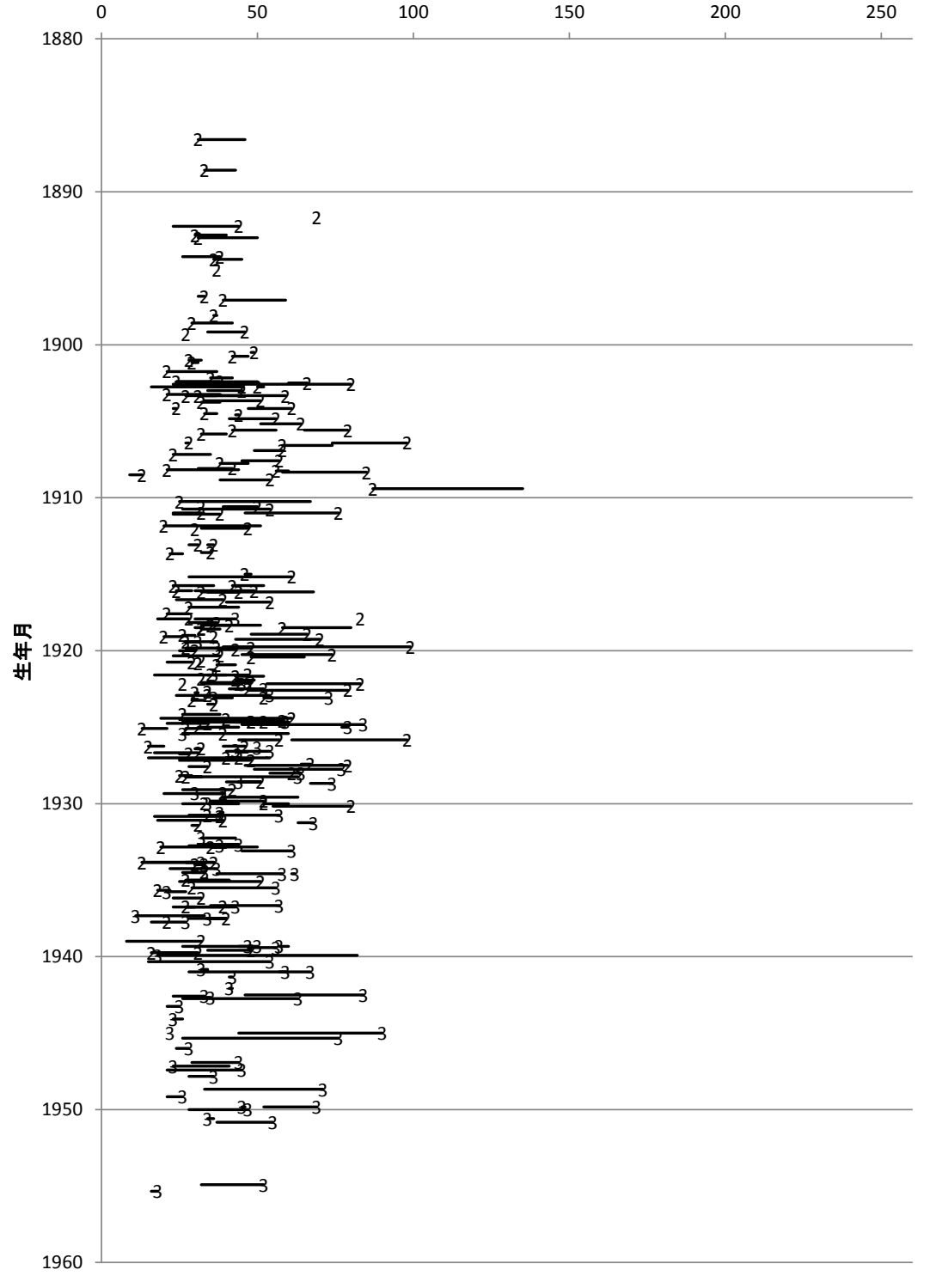
文長



107 議事堂

108

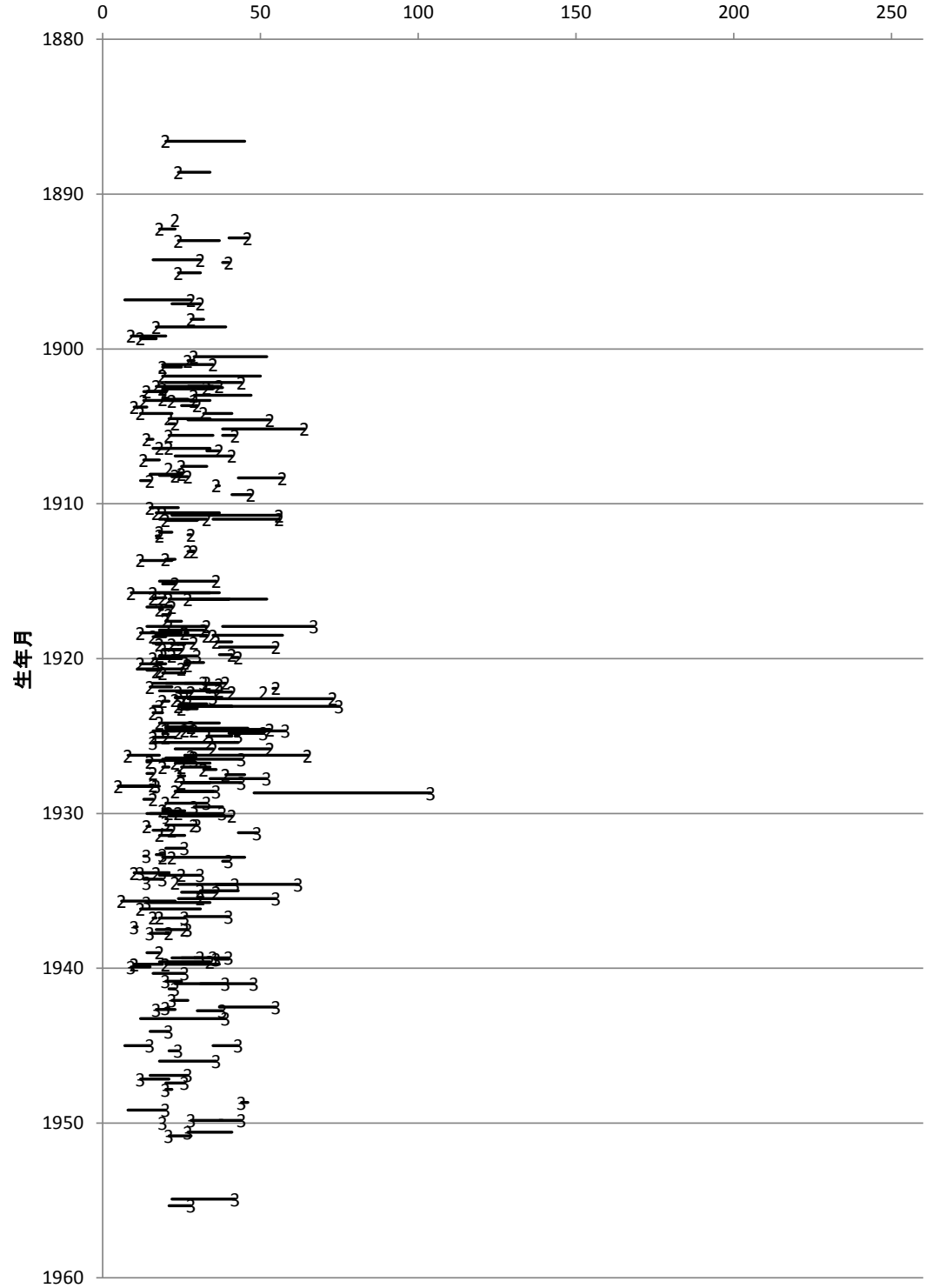
文長



108 医者

109

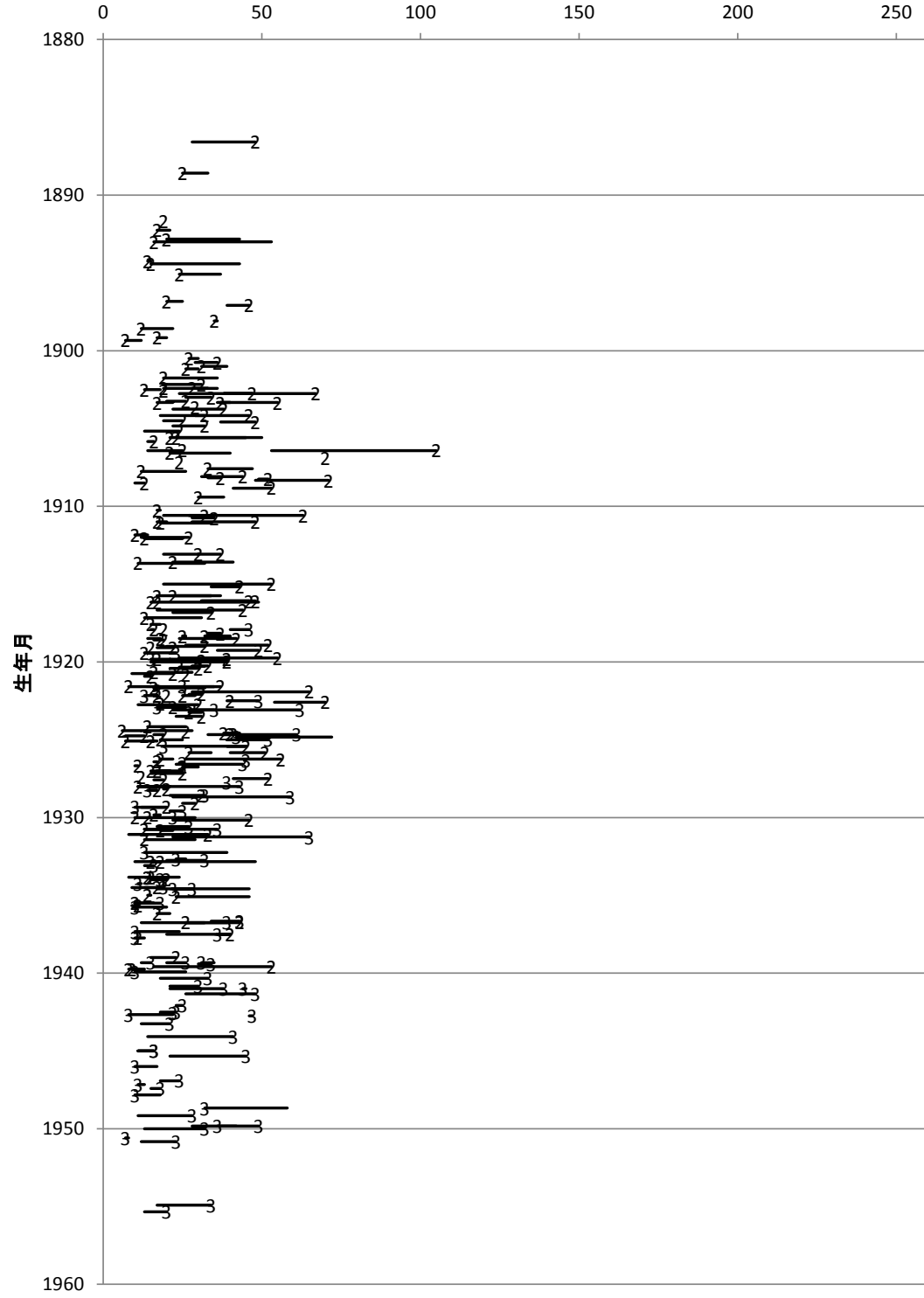
文長



109 席譲られ

110

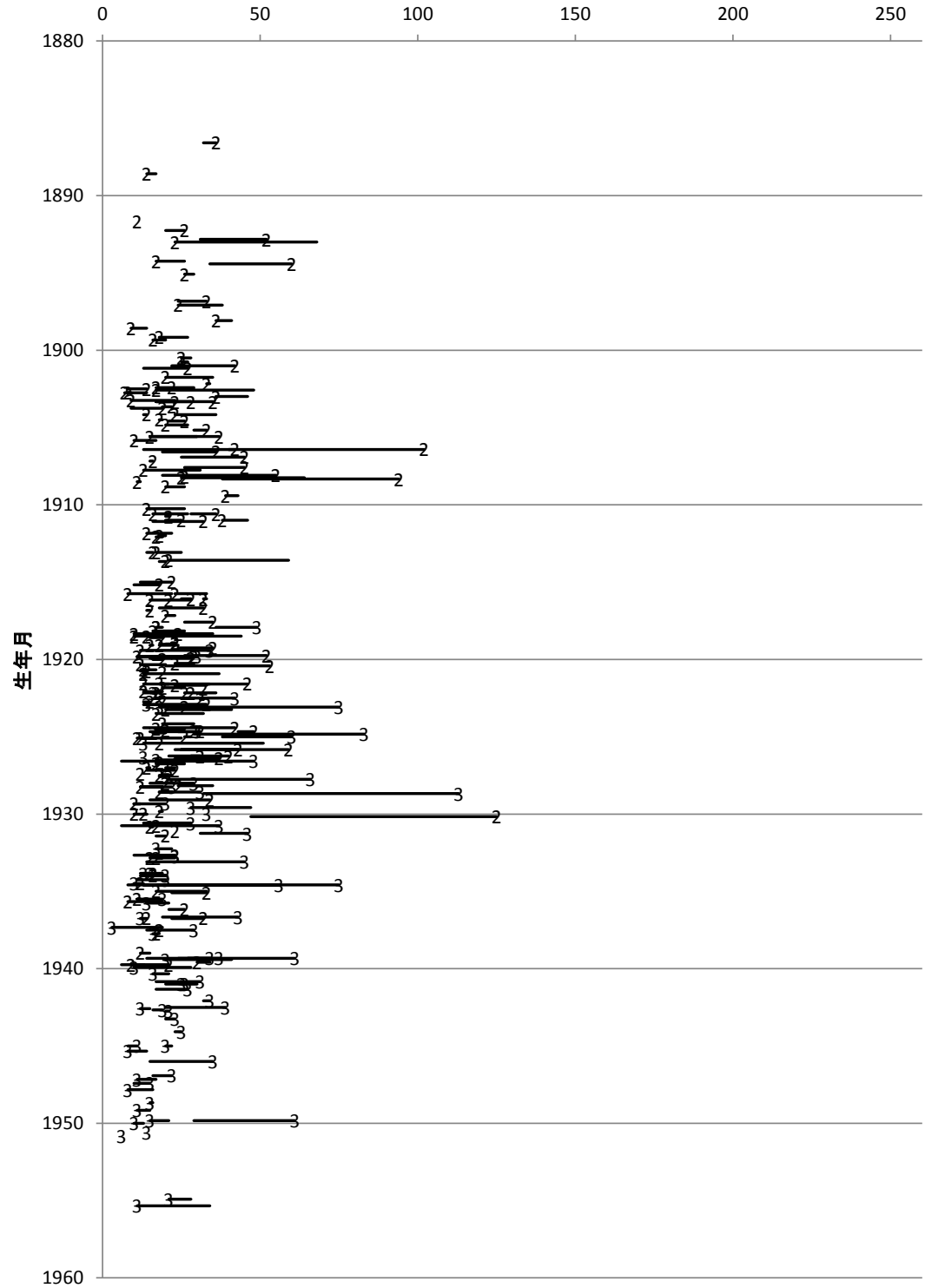
文長



110 おつり

111

文長



111 傘貸し

6. 結論 現代の言語変化

パネルサンプルにおける3次にわたる調査の変化を考察するために、従来の平均値を計算する手法などに代えて、個人の数値を散布図で表現する技法を適用した。回答文が長くなるという経年変化が再確認できた。世の中の状況が変わったために長く話す必要が生じた可能性もあり、言語的に長くなる表現が普及したためもある。また対人配慮のための要素が多く使われるようになったとも考えられる。さらに加齢によって長話をするようになったこともありうる。これらの要因の一部の可能性についてはすでに論じた。他の可能性については、さらに具体的データの分析を通じて解明を目指す。

加齢変化には様々なものがありうる。敬語の成人後採用はそのうちの1種にすぎない。1個人の *idelect* の変化にあたるもので、長年の著作・日記などの追跡や異なった時期の録音録画史料によって、書きことば、話しことばの変化を確認できる。また「記憶時間」を利用して、本人の記憶・思い出を内政してもらうことにより、著作や録音のない普通の人の生涯変化も研究できる。年齢という「見かけ(の)時間」、(繰り返し)調査時点による「実時間」以外にも、現代の言語変化を知る手段はある。実践の結果も多い。研究史展望と理論的考察は別稿に譲る。

本稿は井上(2015)の口頭発表を拡充したものである。グラフはすべて柳村による。なお本報告について、岡崎市の方々に大いにお世話になった。厚く御礼申し上げる。本稿前半は韓国日本語学会機関誌に載る論文と一部分重複する。

参考文献

- 朝日祥之(2008)『ニュータウン言葉の形成過程に関する社会言語学的研究』(ひつじ書房)
- 井上史雄(2011)『経済言語学論考』明治書院
- 井上史雄(2014.12.)「昭和の方言 鶴岡と郊外の言語変化」日本語学 32-15 pp.16-24.
- 井上史雄(2015.2)「敬語の成人後採用」『国語研プロジェクトレビュー』5巻3号 pp.98-107
- 井上史雄(2015年3月8日)「大規模経年調査から岡崎敬語調査パネルサンプルの生年順表示」大規模経年調査研究発表会
- 国立国語研究所(1958)『敬語と敬語意識』(秀英出版)
- 国立国語研究所(1983)『敬語と敬語意識—岡崎における20年前との比較—』(三省堂).
- 国立国語研究所(2010)『敬語と敬語意識—愛知県岡崎市における第三次調査—』科学研究費補助金研究成果報告書 第1~4分冊.

岡崎敬語調査資料集 14

Material for Large-Scale, Long-Term Studies of Japanese

岡崎パネル調査の文の長さ

—繰り返しデータの個別表示技法—

Length of Sentences in Panel Surveys of Okazaki

Individual representation of a repetitive survey

(Ver. 1.1)

日本語の大規模経年調査に関する総合的研究

Comprehensive Research

Based on Large-Scale, Long-Term Studies of Japanese

著： 井上史雄 ・ 柳村裕

INOUE Fumio, YANAGIMURA Yu

発行：平成 27 年 4 月 29 日 29 Apr 2015

国立国語研究所

National Institute for Japanese Language and Linguistics

〒190-8561 東京都立川市緑町 10-2 Tel. 042-540-4300 (代)

10-2 Midori-cho, Tachikawa City, Japan 190-8561