

「日本語非母語話者の読解コーパス」作成のための調査方法

1. コーパス作成のための調査方法の概要

データ収集者はここに示す調査方法に従って調査を行う。調査方法を次のような項目にわけて説明する。

2. 調査当日までにする準備

- 2.1 協力者探し
- 2.2 書類の準備
- 2.3 協力者に対する調査の説明
- 2.4 調査をする日時と場所の決定
- 2.5 協力者による読む素材の選定
- 2.6 意味や訳語がわからない語句の下調べ
- 2.7 協力者にとって難しいと思われる箇所の予測
- 2.8 通訳者との打ち合わせ
- 2.9 機器や読む素材などの準備

3. 調査当日にすること

- 3.1 座る位置と機器の配置
- 3.2 調査の流れ
- 3.3 調査の方法

4. 調査をうまく進めるためのポイント

- 4.1 話しやすい雰囲気をつくる
- 4.2 調査の本番前の練習を効果的に行う
- 4.3 読む素材に書かれている内容以外のこともたくさん話してもらう
- 4.4 調査に影響を与えるようなことはしない
- 4.5 辞書やパソコンなどの使用頻度に注意する
- 4.6 練習のための予備調査をする

付属資料

資料1：協力者の方へ

- 資料1-1：協力者の方へ（日本語版）
- 資料1-2：協力者の方へ（英語版）
- 資料1-3：協力者の方へ（中国語版）
- 資料1-4：協力者の方へ（韓国語版）

資料2：通訳者の方へ

- 資料2-1：通訳者の方へ（日本語版）

資料 2－2：通訳者の方へ（英語版）

資料 2－3：通訳者の方へ（中国語版）

資料 2－4：通訳者の方へ（韓国語版）

資料 3：調査の手順

資料 3－1：調査の手順（日本語版）

資料 3－2：調査の手順（英語版）

資料 3－3：調査の手順（中国語版）

資料 3－4：調査の手順（韓国語版）

資料 4：練習用の読む素材

資料 4－1：練習用の読む素材（やさしいもの）

資料 4－2：練習用の読む素材（少し難しいもの）

資料 4－3：練習用の読む素材（難しいもの）

書類

書類 1：調査協力承諾書

書類 1－1：調査協力承諾書（日本語版）

書類 1－2：調査協力承諾書（英語版）

書類 1－3：調査協力承諾書（中国語版）

書類 1－4：調査協力承諾書（韓国語版）

書類 2：背景調査票

書類 2－1：背景調査票（日本語版）

書類 2－2：背景調査票（英語版）

書類 2－3：背景調査票（中国語版）

書類 2－4：背景調査票（韓国語版）

調査方法説明ビデオ

調査方法説明ビデオ：調査に協力してくださる方へ

2. 調査当日までにする準備

2.1 協力者探し

日本語非母語話者の協力者を探す。協力者は自身の理解過程を分析的に話せそうな人が望ましい。

2.2 書類の準備

次のような書類を準備する。

(1) 調査協力承諾書（書類 1） 2 通

調査協力承諾書には日本語版、英語版、中国語版、韓国語版がある。協力者に応じて、そのどれかを使う。

(2) 背景調査票（書類2） 1通

背景調査票には日本語版，英語版，中国語版，韓国語版がある。協力者に応じて，そのどれかを使う。

2.3 協力者に対する調査の説明

次のようなことについて協力者に説明をする。説明は直接会って行っても，メールなどで行ってもよい。

(1) 調査の内容

「協力者の方へ」（資料1）を協力者に見せて，調査の内容について説明する。

(2) 調査協力承諾書への署名

「調査協力承諾書」（書類1）を協力者に読んでもらう。承諾した場合，調査協力承諾書2通に協力者が署名をする。協力者には署名したものを調査当日に持ってきてもらう。それができない場合には，調査当日に署名してもらう。

署名をした1通は協力者に渡す。ただし，協力者が調査協力承諾書の所有を希望しない場合には，署名は1通だけでよい。

(3) 読む素材

協力者には調査当日に初めて素材を読んでもらうので，事前に素材を読まないように伝えておく。

(4) 辞書やパソコンなどの持参

普段使用している辞書やパソコン，スマートフォンなどを調査当日に持参するように協力者に伝えておく。

(5) 背景調査票への記入

「背景調査票」（書類2）を調査日より前に協力者に送り，記入しておいてもらう。それができない場合には，調査当日に記入してもらう。

2.4 調査をする日時と場所の決定

データ収集者は協力者と相談して，調査をする日時と場所を決める。調査のときの発話は IC レコーダーで録音するため，エアコンの音や，壁による音の反響などに配慮しつつ，できるだけ雑音が少ない静かな部屋を選ぶ。

可能であればデータ収集者は協力者と調査日より前に顔合わせをし，信頼関係を作っておくと，調査がやりやすくなる。

2.5 協力者による読む素材の選定

読む素材は（1）から（3）のすべての条件を満たすものを選んでもらう。

(1) 協力者が読みたいものや読む必要のあるもの

協力者に読みたいものや読む必要のあるものを選んでもらう。データ収集者

が与えるのではない。協力者には、非母語話者向けに書かれた日本語教科書などの日本語学習用教材は選ばないようにしてもらおう。

協力者が自分で適切な読む素材を見つけるのが難しい場合は、データ収集者が助言してもよい。

(2) 協力者の読解能力にふさわしいもの

読む素材は協力者の読解能力にふさわしいものを選んでもらう。

難しすぎると、「わかりません」のような発話ばかりになり、理解過程が把握できなくなる。反対に簡単すぎると、すべてを瞬時に理解するため、知らない語句の意味をどのように推測したかや、解釈がどのように変わっていったかなどについての発話が少なくなる。発話は読む素材の文章をそのまま翻訳したようなものになり、理解課程が把握できなくなる。

(3) ウェブサイトでの長期間の公開に支障がないもの

この読解コーパスは、素材とそれを読んだときの発話を並べて示すものであり、今後、長期間ウェブサイトで公開される。そのため、読む素材はウェブサイトで長期間公開しても支障がなさそうなものを選ぶ。たとえば、著作権フリーのものや、ウェブサイトで長期間閲覧できそうなものを選ぶ。

2.6 意味や訳語がわからない語句の下調べ

データ収集者は素材をよく読み、語句の意味や書かれている内容について不明な点があれば調べ、素材の内容の理解に努める。

通訳者が参加しない調査の場合、調査で使用する言語で何と言うかわからない語句などがあれば、訳語を調べておく。

2.7 協力者にとって難しいと思われる箇所の予測

協力者にとって理解するのが難しいと思われる箇所を予測する。その箇所を協力者がどのように理解するかを予測した上で、協力者が理解した内容を確認するためにはどのように質問すればよいかを考えておく。

たとえば、次のような箇所で理解するのが難しいと考えられる。

(1) 難しそうな語句

読む素材の例：日本において流通革命の必要性が唱えられてすでに四半世紀がすぎた。

予測される発話の例：「4つの半世紀が過ぎました」

予測される発話に対する質問の例：「「4つの半世紀」というのはどういう意味ですか」

(2) 長い文の構造

読む素材の例：小売業地域の構造への影響について、林は首都圏を、山田は関西大都市圏を、佐藤は名古屋都市圏を研究対象として選択している。

予測される発話の例：「林は首都圏。山田は関西大都市圏。佐藤は名古屋都市圏を研究対象として選んでいます」

予測される発話に対する質問の例：「林は首都圏をどうしますか」

(3) 省略された主語

読む素材の例：点検によりガス設備に不具合等が発見された場合には、改善方法などをご説明させていただきます。

予測される発話の例：「点検によりガス設備に問題が発見された場合、改善方法を説明します」

予測される発話に対する質問の例：「だれが説明しますか」

(4) 文脈の意味

読む素材の例：金融危機の要因については、近年多くの研究がなされ、様々なことが解明されつつある。しかし、金融機関と住宅ローンの関係については、これまで十分な考察がなされてきたとは言いがたい。

予測される発話の例：「金融危機の要因については多くの研究がされています。金融機関と住宅ローンの関係についての研究もあります」

予測される発話に対する質問の例：「この部分で筆者は先行研究に問題点があると述べていますか。述べていませんか」

2.8 通訳者との打ち合わせ

データ収集者が協力者の母語を十分に運用できない場合には、通訳者を手配して参加してもらう。通訳者は会話能力だけでなく、日本語の文章を正確に理解できる能力がある人を選ぶ。同時通訳に近い形で通訳できる人が望ましい。

通訳者が参加する調査の場合、次のようなことについて打ち合わせをする。打ち合わせは直接会って行っても、メールなどで行ってもよい。はじめの説明のときに、通訳者に読む素材を送る。

(1) 調査の目的や方法の説明

通訳者に「通訳者の方へ」(資料2)を渡してその内容を説明し、調査の目的や方法、通訳するときの注意点をよく理解してもらう。

(2) どのような質問を考えているかの説明

データ収集者が質問をする可能性のある箇所について、質問をする内容を通訳者に説明する。

(3) 通訳者がする準備の説明

通訳者には素材をよく読んでもらい、通訳する言語で何と云うかわからない語句などがあれば、調査当日までに調べておいてもらう。

2.9 機器や読む素材などの準備

次のような準備をする。

(1) 機器の準備

準備する機器は次のようなものである。

IC レコーダー：1 台（可能なら 2 台）

ピンマイク（モノラルマイク）：2 本

2 本のピンマイクを接続するためのプラグアダプター（例：SONY プラグアダプターPC-239S）：1 個

IC レコーダー用の予備の電池：数個

ビデオカメラと三脚：1 式（協力者が素材のどこを読んでいるか録画する必要がある場合）

協力者に調査方法説明ビデオを見せるためのパソコンやタブレット：1 台

Wi-Fi ルーター：1 個（協力者がインターネットに接続できない環境にパソコンを持参する場合）

調査のときの発話は IC レコーダーで録音する。小さいつぶやきなども明瞭に録音できるようにピンマイクを使用する。協力者とデータ収集者の 2 人の音声を 1 台の IC レコーダーに録音するために、2 本のピンマイクをプラグアダプターに接続する。そのプラグアダプターを IC レコーダーのマイク端子に接続する。

予期しないトラブルが生じた場合でも録音ができているように、可能なら IC レコーダーを 2 台準備する。

協力者が素材のどこを読んでいるかを録画する必要がある場合は、ビデオカメラと三脚も準備する。

協力者がパソコンを持参するときには、そのパソコンがインターネットに繋がらない場合にも対処できるように、Wi-Fi ルーターを準備する。

(2) 協力者が練習のときに使用する読む素材

データ収集者は、協力者が本番前の練習のときに使用する読む素材として、実際に調査で使用するものと同じジャンルで、同じぐらいのレベルの読む素材を準備する。この調査方法に付いている練習用の読む素材（資料 4）を使ってもよい。

(3) 協力者が調査で読む素材

データ収集者は、協力者のために読む素材を 1 部準備する。これとは別に、何かあったときのために、2 部ほど余分に用意しておくもよい。

(4) 飲み物などの準備

休憩時間に気分転換をしてもらうために、可能なら、飲み物やお菓子などを準備する。

3. 調査当日にすること

3.1 座る位置と機器の配置

(1) 座る位置

データ収集者は協力者の隣に並んで座る。そのとき、協力者が読んでいる素材を指し示すことができる距離に座る。通訳者が参加する場合には、協力者を真ん中に挟むようにしてデータ収集者、協力者、通訳者の3人が並んで座り、それぞれが素材を手元に置いておく。

(2) ピンマイクの配置

通訳者が参加しない調査の場合は、2本のピンマイクのうち、1本を協力者の口の近くにつけ、もう1本をデータ収集者の口の近くにつける。

通訳者が参加する調査の場合は、2本のピンマイクのうち、1本を協力者の口の近くにつけ、もう1本を通訳者の口の近くにつける。この場合、データ収集者にはピンマイクがつかないが、データ収集者が大きめの声で話すことで2本のピンマイクに音声録音される。

(3) ICレコーダーの配置

ピンマイクを付けたICレコーダーは調査の邪魔にならないような場所に配置する。ICレコーダーが2台あるときには、予期しないトラブルが生じた場合でも録音ができているように、ピンマイクが付いていないほうのICレコーダーを、すべての発話が録音できるようなところに配置する。

(4) ビデオカメラの配置

ビデオカメラで録画する必要がある場合は、協力者が素材のどの部分を読んでいるかをはっきり録画できる位置にビデオカメラを配置する。たとえば、パソコンでウェブサイトにある素材を読むとき、ウェブサイトのページを次々と移動しながら読むことも考えられる。そのような場合、ビデオカメラでパソコンの画面を録画して、ページの移動の状況がわかるようにする。後で音声データと録画データの照合ができるようにビデオカメラでも音声を録音しておく。

また、ビデオカメラによる録画を確実にを行うために、デュアルディスプレイの設定をするという方法もある。この設定は協力者がウェブサイトを見ているパソコンにもうひとつデータ収集者用のディスプレイをつなぎ、協力者が見ている画面と同じものをデータ収集者のディスプレイでも見られるようにするものである。データ収集者のディスプレイの画面をビデオカメラで録画すれば、協力者の邪魔をすることなく、協力者がウェブサイトのどこを読んでいるか、また、何を調べているかを記録できるので便利である。

3.2 調査の流れ

調査の流れは次のとおりである。

(1) 「調査協力承諾書」(書類1)と「背景調査票」(書類2)を回収する

事前に協力者に渡してあった「調査協力承諾書」(書類1)と「背景調査票」(書類2)を回収する。署名や記入がまだされていないときには、その場でしてもらう。

(2) 調査の目的や手順を説明する

協力者に「調査の前の説明」(資料3)を渡し、目を通してもらう。「調査の前の説明」には日本語版、英語版、中国語版、韓国語版がある。協力者に応じて、そのどれかを使う。データ収集者は「調査の前の説明」を読み上げ、協力者に調査の目的や手順を説明する。

(3) 調査方法説明ビデオを見てもらう

協力者が調査方法説明ビデオを事前に見ていない場合は、調査方法説明ビデオを見てもらい、読みながらどのように発話するかを理解してもらう。

(4) 練習してもらう

協力者に調査に慣れてもらうために、練習用の読む素材を使って初めに練習してもらう。練習の段階から調査のすべての発話をICレコーダーで録音する。ビデオカメラで録画する必要がある場合は録画する。

データ収集者は協力者の読み方や発話の仕方が適切に行われるように助言する。協力者が発話することに慣れたら、調査を開始する。

(5) 調査を開始する

素材を普段読むときと同じように読んでもらう。普段、辞書やパソコン、スマートフォンなどを使用しているなら、この調査でも使用してもらう。素材は前から順に読む必要はなく、どのような順序で読むかは普段どおりにしてもらう。協力者が読む必要がないと判断した部分があれば、そこは読み飛ばしてもらってもよいが、そのとき、読み飛ばす理由を協力者に確認する。

(6) 休憩を挟む

協力者や通訳者がそれほど疲れているように見えなくても、集中力がとぎれないように1時間ほど調査をしたら休憩を入れるようにする。心理的にも肉体的にも負担をかけないようにして、質のよいデータをとることを心がける。

(7) 読み終わった後の感想を話してもらう

可能なら、協力者に読み終わった後の感想を話してもらう。

(8) 調査で読んだ素材を回収する

調査中に協力者が素材に書き込んだメモなどは、調査後にデータの文字化や分析をする際の貴重な参考資料となるので、調査が終わったら読んだ素材を回収する。

3.3 調査の方法

協力者に素材を読んでもらい、それと同時に理解過程を、母語あるいは母語に準じる言語で少しずつ話してもらう。データ収集者は協力者に読む素材の内容を確認する質問をし、答えてもらう。

調査は次のように行う。

(1) 素材を読んでもらい理解過程を話してもらう

話してもらうことは、次のようなことである。

a. 読む素材に書いてある内容

読む素材に書いてある内容について理解したことを、たとえば、次のように話してもらう。

読む素材の例：ガス漏れの危険がある場合は、恐れ入りますが下記までご連絡をお願いいたします。

発話の例：「ガスは危険で怖いですから連絡をお願いします」

b. 頭の中でどのような推測をしているか

たとえば、知らない語句があったとき、その意味を、文脈を手がかりにして、どのように推測したかについて話してもらう。

読む素材の例：点検によりガス設備に不都合が発見された場合には、係員より改善方法などをご説明させていただきます。

発話の例：「不」から始まる漢字3つのことばの意味がわかりません。でも、「不」が付いているから、何か悪いことの意味だと思います。これはガスの点検についての文章ですから、ガスの問題のことかもしれません」

c. 読んでいる箇所について感じたこと

そのとき読んでいる箇所について感じたことを話してもらう。

読む素材の例：ガス事業法に基づき3年に一度、敷地内および屋内のガスの配管やガス風呂がま、ガス湯沸器などの点検にお伺いさせていただきます。

発話の例：「この文は長くて、ちょっと難しいです」

d. わからないことや判断に迷っていること

わからないことや判断に迷っていることなどもそのまま話してもらう。

読む素材の例：お手持ちのガス器具をあらかじめご用意いただくと、大変助かります。

発話の例：「ガス器具を用意するのは、だれでしょうか。私でしょうか。ガス会社の人でしょうか。よくわかりません」

e. 読んでいる箇所について関連して思い出したこと

そのとき読んでいる箇所について関連して思い出したことなどを話してもらう。

読む素材の例：ご訪問の際は、お客さまから点検費用をいただくことはございません。

発話の例：「この「際」という漢字は、たまたま昨日辞書で調べたばかりなので意味がわかります。「～したとき」と同じ意味です」

f. 解釈の変更

既に読んだ部分の意味の解釈が変わったときには、読み進めるにつれてどのように変わってきたかについて話してもらおう。

読む素材の例：日時変更をご希望される場合は、恐れ入りますが下記までご連絡をお願いします。

発話の例：「さっきは「恐れ入りますが」を「怖い」と言いましたが、それは違っていました。今読んだこの文には「日時変更を希望」と書いてあるので、「怖い」と関係がないと思います」

理解したことを話してもらうときには、読む素材の1文単位で言ってもらえば必要はなく、発話しやすい語句レベルに区切って言ってもらってもよい。

協力者が沈黙している時間が長いとき、「どこを読んでいますか」「何を考えていますか」「どこが難しいですか」などと話しかけるようにして、頭の中で考えていることをいろいろ話してもらおう。

(2) 辞書などを使って読む場合には、その状況を話してもらおう

協力者には辞書などを普段読むときと同じように使ってもらって読んでもらう。どのような辞書を使って何を調べようとしているかという自分の行動や、調べているときに頭の中で考えていることなどをすべて話してもらおう。たとえば、次のように話してもらおう。

読む素材の例：敷地内および屋内のガスの配管やガス風呂がま、ガス湯沸器などの点検にお伺いさせていただいております。

発話の例：「ガス風呂」の次のことば「がま」ということばがわからないので、スマートフォンの辞書で調べます。「goo 辞書」の日中辞書を使います。「がま」と入力すると意味が3つ出てきました。1つめは「植物のがま」で、2つめは「ガマガエル」、3つめは「がまんする」です。辞書からは「ガス風呂」と合う意味が見つかりません。ここは、とりあえず飛ばします」

(3) 辞書やパソコンの使用状況を記録する

協力者が読みながら辞書やパソコンなどを使用した場合、どのような辞書やウェブサイトを使ったかをデータ収集者が記録する。どの語を調べようとして、どのような語を入力したら、どのような意味が候補として出てきて、そこからどのような過程を経て最終的にどう理解したかも記録する。これらの記録はメモをとるかカメラで撮影するなどして行う。

辞書やパソコンなどの使用状況が、協力者の発話だけではわからないときは、質問して確認する。

(4) 質問に答えてもらう

協力者の発話だけではどのような意味に理解したかがよくわからないときや、協力者が言及しなかった点があるとき、それを確認するために協力者に質問をし、答えてもらう。協力者は理解したことすべてを話すわけではないため、

データ収集者は協力者の発話をよく聞いて柔軟に対応する。

読む素材の中で協力者にとって難しいと思われる箇所を事前に予測しておくが、それについての質問は協力者の反応を見ながら、その場で臨機応変に行う。質問は協力者の思考を妨げないように状況をよく見て行う。

データ収集者が読む素材の中の語句に言及するときに、その語句を日本語で発音すると、協力者の理解の助けになることがある。そのようなことを避けるために、語句を発音せずに指で指し示すようにする。

協力者が自分の理解したことを説明したときに、「なぜそう思いましたか」のように質問をして、協力者の理解過程を明らかにするように心がける。協力者が間違ったことを言ったときだけではなく、難しいと思われた箇所について正しく理解できたときにもそのように質問する。

(5) 必要に応じて、読む素材のどの部分についての発話であるかを記録する

たとえば、協力者が読みながら、「このことばはわかりません」と言う場合、調査の後で録音データを聞いたときに、「このことば」が読む素材の中のどの語を指しているかが不明確になる。それを避けるために、データ収集者は調査のときに必要に応じて、「このことば」というのは、2 ページ、10 行目のカタカナのことばですね」などと発話して録音しておく。

(6) 通訳者にデータ収集者と協力者の発話をすべて通訳してもらう

通訳者が参加する調査の場合、発話をすべて同時通訳に近い形で通訳してもらう。

4 調査をうまく進めるためのポイント

4.1 話しやすい雰囲気をつくる

調査はデータ収集者と協力者が基本的に1対1で行うため、友好的な関係が築けるように心がける。調査の本番前に行う練習のときには、「それでいいですよ」「上手ですね」のように、肯定的で勇気づけることばをかけるようにし、安心して発話してもらう環境を作る。

4.2 調査の本番前の練習を効果的に行う

協力者にとって素材を読みながら同時に理解したことを言語化するのは慣れるまで難しいこともある。そのため、調査の前の練習は大切である。協力者がする発話のモデルを示すときには、調査方法説明ビデオを活用する。

4.3 読む素材に書かれている内容以外のこともたくさん話してもらう

協力者が読む素材に書かれている内容を正確に翻訳することだけに集中して、翻訳以外の発話、たとえば、推測していることやわからないことなどをほとんど言わないことがある。そのようなときには、「この調査は正確に翻訳することが目的ではあり

ません。〇〇さん（協力者の名前）の理解過程をそのまま知りたいと思っています。ですから、迷ったり、推測したりしていることをそのまま話してください」のように声をかける。

4.4 調査に影響を与えるようなことはしない

データ収集者の発言や行動が調査に影響を与えないように細心の注意を払う。たとえば、質問の答えのヒントになることを言ったり、協力者が間違ったことを言ったときに、間違ったと気づかせるような反応をしたりしない。

4.5 辞書やパソコンなどの使用頻度に注意する

協力者に辞書やパソコンなどの使い方については事前に、「普段読むときと同じように使う」ことをしっかり伝えておく。普段の読み方とは違うことはしないようにしてもらう。

たとえば、協力者が辞書やパソコンなどを使用する頻度が多いように感じたら、「いつも読むときにそんなによく辞書を使いますか」のように質問をする。いつも辞書を使わないようであれば、調査のときにも使わないようにしてもらう。

4.6 練習のための予備調査をする

データ収集者は、コーパスとして使えるようなよいデータを収集するために、調査方法を十分把握した上で、練習のための予備調査をある程度行うとよい。