I-JASデータの社会文化的考察

トムソン木下千尋*

1. はじめに

本稿は、多言語母語の日本語学習者横断コーパス (I-JAS, International Corpus of Japanese as a Second Language) を社会文化的視点で考察するものである。I-JASは、国立国語研究所が中心となり、12の言語背景を持つ日本語学習者と日本語母語話者の話し言葉データ、及び書き言葉データを集めた言語コーパスである。第二言語としての日本語のコーパスとしては最大規模で、学習者1000名、日本語母語話者50名のデータが順次公開されてきている¹。一方、社会文化的アプローチの視点は、人の営みは何らかの社会文化的な文脈の中で起こるとし、言語も社会文化的な相互行為を通じて構築されると捉える (佐伯1996)。コーパス研究の基本的な考え方は、データから何らかの普遍性を探ろうとするものであろうが、社会文化的アプローチの考え方は、本質的な『日本語』、あるいは“the Japanese”の存在を疑問視している。

本研究では、I-JASの英語母語学習者データと中国語母語学習者データ、さらに日本語母語話者データを使い、I-JASデータとはどんなものなのか、つまり、I-JASデータはどのような社会文化的文脈に埋め込まれているのか、さらに、データ収集のプロセスでは何が起こっていたのか、すなわち、どのような社会的相互行為により、どのような言語がデータ収集を媒介したのかを考察する。

2. 社会文化的な視点から見た言語

ヴィゴツキーに端を発する社会文化的アプローチは、学習は社会から孤立した個人の営みであるという従来の捉え方から離れ、学習は社会、文化、歴史的な相互行為の中で文化的な道具 (たとえば、ことば) に媒介された活動から生まれると考える (石黒2004)。言語も同様に、社会的、文化的、歴史的視点で捉えられる。Maher (2017) は、「ことば」は、過去の歴史や、話者の家族、友人のことばを踏襲し、その人の故郷の、何世代もの話しことばも含めたモザイクである (訳責筆者) と言う。すなわち、例えば、私が今、この特定の場面で、様々な要因に影響されながら書いている「ことば」は、本質的な『日本語』、あるいは“The Japanese”とは別のものである。

3. I-JASデータとは

I-JASプロジェクト²は、国立国語研究所のホームページに依ると、その目標がI-JASデータの構築により、第二言語習得や日本語教育の研究を行うことである。従って、I-JASデータの収集の目的は第二言語習得研究や日本語教育に役立つデータを集めることということになる³。

I-JASデータ収集の過程は、以下の通りである。まず、20カ国から研究者が集まりプロジェクトチームが結成された。そこで、どんなデータを集めるか、どのように集めるのかを検討し、データ収集者 (調査者) のトレーニングが行われた。調

*ニューサウスウェールズ大学教授

査者はみな大学などで研究をしているほとんどが中年以上の日本語母語話者教師である。各国では参加者（学習者）が募集され、調査者が各国に向いた。各地で対面でデータ収集が行われ、録音された。データの書き起こし、タグ付けは現在も行われていて、データは段階的に公表されている。

I-JASのデータには、話しことばデータ、書きことばデータと学習者の日本語能力レベル判定テストの結果が含まれている。学習者の性別、母語などを示すフェイスシートもある⁴。本稿では、話しことばデータの一部である「対話」データを見て行くこととする。

4. コーパス

ここで、コーパスというものについて少し述べておきたい。コーパス研究は英語におけるものが盛んで、有名で大きなものには、British National Corpus、Cambridge and Nottingham Corpus of Discourse in Englishなどがあり、オーストラリアにも Australian National Corpus of Englishがある。この中にはラジオ放送の録音、インタビューの録音、手紙、小説、詩歌など、実際に使われた言語が収集されている。日本語のものも、現代日本語書き言葉均等コーパス、日本語話しことばコーパスなどがあり、この両者は国立国語研究所からアクセスできる。

コーパスは「実際に使用された言語の産出データを特定の目的及び方法を持って収集・電子化したもの」（投野2006）と定義されている。研究者にとってはありがたいデータであるが、その限界も指摘されてきている。古くはChomsky（1962 in Tono 2005）がコーパスは不完全で偏っていると指摘している。先のラジオの録音などは、既に行われた言語行動の記録だが、データとして使われることが分かっていて収集されたものは、参加者が録音されることを知っているため社会的なものや当たり障りないものに偏る（Carter 2004）こ

とも指摘されている。I-JASデータのように研究のために収集されたコーパス・データは、現実社会におけるデータ収集過程という特殊な文脈の中に位置づけられていると言える。

5. I-JASデータの検討

ここではI-JASデータの中から調査者と学習者の間の「対話データ」を取り上げる。ホームページの情報に依るとこの対話データは「学習者と調査実施者が自然な会話を30分程度行う」と記述されている。I-JASの一回目に公開されたデータの中からオーストラリア人英語話者（EAU）3名、中国の中国語母語話者（CCM）3名、日本語母語話者（JJJ）3名を無作為に抽出し検討対象とする。

対話データは、調査者が以下のトピックを提示しながら、参加者をリードして行く形で行われた会話の書き起こしである。対話はOPI（Oral Proficiency Interview）に類似した形で進行したが、OPIとの違いは、評価が関わらないことだ。

- ・ どうして日本語を勉強しているのか。
- ・ 自分の町で有名なものは何か。
- ・ 誕生日には大抵なにをするか。
- ・ 子どものころどんな子どもだったか。
- ・ 将来何がしたいか。
- ・ お金と時間、あるいは、田舎と都会でどちらがいいか。

5-1. 質問—応答

対話データにはいくつか特徴が見られた。まず、「質問—応答の形」の多用である。

抜粋1

- ①調査者：すごく今までで怖かったこととか一恐ろしかったことってある？
- ②CCM52：恐ろし（恐ろしい）ことは一、恐ろしい物はありません（あります）{笑}
- ③調査者：あ何（なん）ですか？

④CCM52：あーへび(蛇)〈{笑}〉、はいへび(蛇)
はとも怖いです {笑}

⑤調査者：見たことがありますか？

抜粋 1 は調査者と中国語母語話者CCM52のデータである。調査者が質問し、CCM52が応答する形で話が進んでいる。①で調査者が質問形式で新トピックを提示し、②でCCM52が応答しているが、それだけで、展開はない。③で調査者が関連質問をし、それに対して④でCCM52が応答する。ここにも応答以外の発展は見られない。⑤では、調査者がさらに関連質問をしている。

ここに見られるような「質問—応答の形」がデータの大部分を占めていて、学習者側からの反復要求や意味確認などはほとんど見られなかった。6名の学習者の対話、1848ターンの中、学習者が調査者に質問したのは以下の2例のみであった。抜粋 2

①調査者：お金と時間、どちらがたくさんあった方(ほう)がいいですか？

②CCM14：え？

③調査者：どっちがたくさんほしいです

④CCM14：お金は、何(なに)？

抜粋 3

①CCM52：本が好き、子どもの一時は『ハレーポーター(ハリポーター)』

②調査者：ハリ、ハリポ

③CCM52：ご存じで

④調査者：もちろん

これらの「質問—応答の形」の対話は教室によく見られるIRF(質問Initiation-応答Response-フィードバックFeedback)(例えば、Waring 2009)のパターンを踏襲している。このような教室ディスコースに慣れ親しんだ学習者は、初対面で年上の母語話者である調査者との一対一の対話で、この教室ディスコースをよりどころとしたのだろう。

5-2. 新トピックの提示

日本人同士の対話の場合、調査する側、される側に関わらず対話に新しいトピックを随時導入しているが、学習者との対話では、調査者が質問の形でトピックを提示することが圧倒的で、学習者が自分から新トピックを提示することはほとんどなく、まるでインタビューようになっていた。抜粋 4

①EAU03：(前略) ..二〇一〇年は日本で交換留学に行きました

②調査者：あ、どこに行きました？

③EAU03：えっと岐阜県のぎぶ市(岐阜市)に行きました(へえ) あのかげつ(九か月)くらい

抜粋 4 はオーストラリア人英語話者EAU03と調査者のデータである。交換留学の話をしていて、調査者が②で留学先の場所を聞くと、EAU03が③で場所についての返答をした後、下線の部分で留学の滞在期間という新しいトピックを提示している。学習者の方から新しいトピックを提示したのはこの1例のみであった。

日本人同士だと、どちらからもトピックが出されることを考えると、調査という文脈で、調査する人とされる人の役割分担が明確化したために学習者が受け身になったと考えるのは難しい。むしろ、前節で明らかになったように、学習者は調査者に「先生」という役割を想定するしかなかったのではないだろうか。多くの学習者の場合、年上の日本人との関係は「先生」との関係で、学習者には他のレパトリーがない。後述するが、実際、学習者は調査者を「先生」と呼んでいる。擬似的な教師—学生という役割分担の中で、学習者は自分からトピックを提示することができなかったと考える方が妥当であろう。

5-3. 会話の共構築と専有

日本語会話の特徴として、会話の共構築はよく取り上げられるが(例えば、水谷1983の「共話」、

日本人参加者間の対話には、以下の抜粋5のように調査者、非調査者の役割に関係なく、お互いにトピックを出し合い、ストーリーを一緒に作って行く傾向が見られた。

抜粋5

- ①JJJ03：はい、蟹とか食べづらいし、美味しくないしもう全然よくわかんない（はーはーはーはー）って思って
- ②調査者：へー（はい）で、こっちに来て、味が分かって
- ③JJJ03：あー案外美味しいじゃんみたいな {笑}
- ④調査者：んーで鳥取の良さを見直すっていう
学習者と調査者の場合は対話を通じて共構築は行われていたものの、その形は日本人同士のものとは異なっていた。

抜粋6

- ①調査者：寮に住んでいますか？
- ②CCM14：料理？
- ③調査者：寮、寮に住んでいますか？どこに住んでいますか？
- ④CCM14：寮？あはい、えっとりよ寮に住んでいます

抜粋6では②でCCM14が「料理？」と言ったのを聞いて、調査者はCCM14が「寮に」と「料理」を誤解したと気づき、③で「寮」と明確に発音し、質問を反復し、さらに「寮」が分からなくても答えられるような質問を追加している。これを聞いてCCM14は、「寮」という単語を思い起こし、④で確認し、正しく返答している。このように、学習者との対話の場合は調査者が主導的に意味交渉を繰り返す形での共構築となっていた。

又、この共構築の過程では学習者がことばを「専有する」ということが起こっていたとも解釈できる。「専有 (appropriation)」とは、ある人が社会に存在する心理的道具（例えば、ことば）を新しい場面で採用し自分のものにする（Leontiev 1981）である。

抜粋7

- ①調査者：えっとCCM14さんは今日、どうやってここまで来ましたか？
- ②CCM14：うーん、調査を受けました
- ③調査者：どうやって来ましたか？歩いて？
- ④CCM14：あー、はい
- ⑤調査者：歩いてきましたか？
- ⑥CCM14：歩いてきました

抜粋7では、調査者の①の質問を聞いたCCM14が②で「どうやって」と「どうして」を混同してしまう。③で調査者は混同に気づき、質問を反復し、さらに例をスキヤフォールディングとして提示している。④でCCM14は「はい」と言っているが、調査者はまだ不安だったのだろう、⑤で確認している。それを受けて、CCM14は⑥で確信のある、文法的にも正しい返答をしている。ここでは何度か調査者主導の意味交渉が繰り返されているが、最終的に産出された「歩いてきました」という文は、初級で必ず学習される句型で、しかも「歩く」は基本語彙なので、CCM14が教科書的知識としては、知っていたと推測される。しかし、その知識は記号的なものであって、CCM14にとって特別な意味があったわけではないだろう。それが、この対話で意味交渉をしたことで、「歩いてきました」はCCM14が調査当日に実際に歩いてきた体験とつながり、人込みをかき分けて歩いてきたとか、遅れそうになって急いで歩いてきたとか、何らかの実際場面での文脈が与えられ、CCM14がこの「歩いてきました」を自分のものとして、専有したと言えるだろう。すなわち、学習者は専有と言う形の学習を体験していた。

もちろん、複数のスキヤフォールディングにも関わらず、失敗に終わった共構築の試みもあった。

5-4. 調査者の知識の再構築

対話という体験から変化したのは学習者の日本語だけでなく、調査者の方にも知識の再構築が起こっていた。

抜粋 8

- ①調査者：そう、それから『セーラームーン』？
 ②EAU03：『セーラームーン』
 ③調査者：おおおお、ええテレビでやってたんですか？

小さい頃見ていたテレビ番組の話だが、EAU03がオーストラリアでセーラームーンというアニメを見ていたという知識は、オーストラリアという土地では新参者である調査者にとっては全く新しいものである。この会話の結果、調査者にとって、セーラームーンは、「日本のテレビアニメ」という意味付けから「オーストラリアでも放送されていた日本のテレビアニメ」という意味付けに変わった。すなわち、知識の再構築が起こったのである。このように、調査者は調査地の様々な事柄について学んでいた。

5-5. 学習者と調査者の対話

以上のように、学習者と調査者の対話は、I-JASのホームページにある「学習者と調査実施者が自然な会話を30分程度行う」という記述から想像されるものとは違うようだ。I-JASのデータは、調査者が常にリードし、話題を決め、質問をし、学習者はそれに答えていた。調査者と学習者の役割と関係が、擬似的な教師と学生の関係に落とし込まれていて、それが変化することはなかった。一般的にいう情報交換のための「自然な会話」というより、教室ディスコースの形をとったデータ収集のためのインタビューになっていたと言える。

6. I-JASデータはどのような社会的文脈に埋め込まれているのか

I-JASデータは、データ収集という文脈に埋め込まれている。このプロセスに参加している調査者は、みな日本人母語話者で大学などに所属する日本語教師歴が長い研究者である。学習者に比べ

て年齢はかなり高い。学習者は大学生で、大学側のデータ収集に協力してほしいという募集に応える形で参加していた。調査の場所も大学内の会議室が主で、非常に正式で大事な活動に参加していると感じただろう。この場で一対一の対面式で対話が行われた。対話はデータ収集のために作られた擬似真性タスクの一つである。この社会的文脈のなかで、対話が教室ディスコースに似た質疑応答の繰り返しになってしまったのは当然のことだろう。

しかし、この状況は、学習者にとって普段体験することのない母語話者と一対一で会話する貴重な機会であったことも事実である。

7. どのような社会文化的相互行為により、どのような言語がデータ収集を媒介したのか

会話は、データ収集という社会的相互行為が生み出したもので、そのデータ収集を媒介していたのは擬似的な教師と学生の関係が生み出す擬似的な教室ディスコースであった。そこでは「教師」が質問をし、スキャフォールディングを与え、新しいトピックを提示していた。I-JASのデータの本来の機能は研究のためにデータを収集することであるから、これはむしろ当然のことと言える。

しかし、データ収集のプロセスを通じて、I-JASのデータ収集には別の社会文化的相互行為が二つ生まれてきていた。一つ目は、データ収集が学習者と調査者の学びのサポートをしていたことである。インタビューと言う相互行為を通じて、学習者が意味記号としての「ことば」を専有し、「自分のことば」としていく様子が見られた。学習者だけでなく、調査者も、会話を媒介として知識の再構築をしていた。

二つ目はデータ収集のプロセスが学習者のモチベーションを上げたことだ。

「ほんの短い時間でしかたけど、会話やロールプレイをして、すごく楽しかったです。先生が思う

とおり、私の日本語が短すぎてバラバラになってしまう場合が多いです。しかし、日本語を話すのが好きなので、もっと頑張りたいと思います。私の話をちゃんと聞いてくださったり、アドバイスをくださったりして、本当にありがとうございます。」(迫田2014, p.136)

これは、学習者がデータ収集後に書いたコメントの一つである。普段日本語母語話者と接する機会のない学習者にとって、まとまった時間を調査者と過ごし、日本語で会話し、タスクを行い、自分の日本語に関してフィードバックをもらう相互行為は貴重で、学習意欲を上げる効果があったと思われる。学習者が調査者を「先生」と呼んでいることにも注目したい。

8. まとめ

本稿ではI-JASの対話データとデータ収集のプロセスを社会文化的視点で考察してきた。

I-JASの対話データは、データ収集のための会話で、一般的な普通の会話という意味での「自然な日本語」ではない。むしろ、教室ディスコースに近く、インタビューのようなものであった。また、I-JASのデータ収集プロセスは、本来の機能である研究のためのデータ収集だけでなく、別の機能にも埋め込まれていた。データ収集のプロセスは、学習者に調査者とコミュニケーションする機会を提供し、日本語学習のモチベーションを上げ、調査者が既存知識を再構築する機会をも与えていた。そこから生まれたI-JASデータは、学習者一人のデータではなく、調査者と学習者の両者によって共構築されたものである。I-JASデータは、調査者と学習者のその時、その場所での相互行為から生まれたテキストの集まりである。つまり、本質的な『日本語』、あるいは、“The Japanese”とは違う性質のものであるといえる。

本稿の検討から、課題が見えてくる。I-JASのデータを使って研究を行う際に、私たちはどのよ

うな立ち位置をとればいいのか。例えば、共構築されたデータを学習者データとして利用していいのか。また、I-JASデータの特質は研究の中でどのように記述されるべきなのだろうか。特殊な擬似的教室ディスコースを使って、何らかの一般化を行っていいのか。

I-JASをはじめとしたコーパスは忙しい研究者の味方であると思うものの、その利用の仕方には様々な課題があることは明らかだ。

参考文献

- 石黒広昭 (編著) (2004) 『社会文化的アプローチの実際』 北大路書房
- 佐伯胖 (1996) 「学びあう共同体をつくる」 佐伯胖・中西新太郎・若狭蔵之介 (編) 『学びの共同体』 青木書店 pp.9-32
- 投野由紀夫 (2006) 『投野由紀夫のコーパス超入門—コーパスでわかる英語学習のコツ』 小学館
- 迫田久美子 (2014) 『海外連携による日本語学習者コーパスの構築 中間報告書』
- 水谷信子 (1983) 「あいづちと応答」 水谷修 (編) 『講座日本語と表現 3 話し言葉の表現』 筑摩書房 pp.37-44
- Carter, R. (2004) *Language and Creativity: The Art of Common Talk*. London: Routledge.
- Leontiev, A.A. (1981) *Psychology and the Language Learning Process*, Oxford: Pergamon.
- Maher, J. (2017) “Language diversity and the languages of Japan.” A keynote presentation at the Biennial Conference of the Japanese Studies Association of Australia, July 2017.
- Tono, Y. (2005) “Corpus-Based SLA Research: State of the Art Learner Corpus Studies.” in Minani, M. et al. (Eds.) *Studies in Language Sciences* (4)45-77.
- Waring, H.Z. (2009) “Moving out of IRF: A single case analysis.” *Language Learning*. 59(4)796-824.