

令和元年6月20日現在

機関番号：32689

研究種目：基盤研究(B) (一般)

研究期間：2015～2018

課題番号：15H03227

研究課題名(和文) A corpus-based multi-level analysis of spoken French produced by pre-advanced Japanese learners of French

研究課題名(英文) A corpus-based multi-level analysis of spoken French produced by pre-advanced Japanese learners of French

研究代表者

Detey Sylvain (DETEYS, Sylvain)

早稲田大学・国際学院・教授

研究者番号：00548927

交付決定額(研究期間全体)：(直接経費) 12,300,000円

研究成果の概要(和文)：本プロジェクトでは、初中級学習者に焦点を当て、フランス語の発話を音韻面だけでなく、様々なレベルから中間言語発達分析を進めてきた。41名の学習者のフランス語発話を録音し、そのうち17名が第1回目の調査から1年後に再び録音を行い、経年調査を行った。合計25時間30分(214,278語)分の自然会話を転写し、75,873のアノテーションを手動で付加した。アノテーションが付加されることで、音韻や語彙文法レベルに関する初中級学習者の発話特徴の一端を明らかにすることができた。このアノテーション付きコーパスを活用することで、初中級レベルの学習者に対して効果的な指導を行うことが可能となるだろう。

研究成果の学術的意義や社会的意義

This project collected speech data from Japanese learners of French with an intermediate level of proficiency to determine what problems remain on the pronunciation level, but also lexical and grammatical levels. This will help French language educators in Japan improve their teaching curricula.

研究成果の概要(英文)：The current project aimed at furthering our knowledge of Japanese-French interlanguage development by building a corpus of spoken French produced by pre-advanced Japanese learners of French and annotating the data at several levels of analysis. The underlying research objective was to provide a first overview of the divergences that remain in the spontaneous speech of pre-advanced Japanese learners using a basic annotation procedure.

41 learners were recorded in four universities, out of which 17 were recorded a second time after one year for longitudinal purposes. A total of 25 hours and 30 minutes (214'278 words) of spontaneous speech were orthographically transcribed, aligned, and manually annotated with ad hoc conventions, resulting in 75'873 annotations. The data has been loaded to a dedicated server with an ad hoc data mining tool to search the corpus. Our annotated corpus will be useful to prioritize teaching topics to avoid lingering divergences at such level of learning.

研究分野：外国語教育

キーワード：外国語教育 応用言語学 コーパス フランス語 日本人学習者

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1 . 研究開始当初の背景

Over the past 15 years, the use of spoken corpora in the field of second language education has been gradually attracting more and more interest in the communities of researchers and teachers alike, despite specific challenges that remain to be solved in tackling non-native speech (vs non-native writing). In the case of Japanese learners of French, only few researchers have compiled such corpora, and the most important one – IPFC-Japanese, to which the current project is attached – was originally designed to focus on pronunciation. In order to expand the scope of inquiry, the current project was launched.

2 . 研究の目的

While previous research based on spoken corpora of Japanese learners of French focused on beginner learners' pronunciation in Tokyo, the current research was aiming at expanding the scope of research in three ways: (a) expanding the scope of analysis (from phonetics to complex lexico-grammatical structures); (b) examining higher levels of proficiency to assess the remaining interlinguistic divergences (B1-C1 according to the levels of the CEFR); (c) including students from several universities (two from Tokyo and two from Fukuoka) to take into account the diversity of learners. The objective of the research was to test the possibility of using a basic annotation procedure to obtain an overall picture of the learners' productions, with a view to use these results both for research-oriented and pedagogy-oriented implications.

3 . 研究の方法

The project method can be described in four parts: 1) recording protocol, 2) annotation procedure, 3) database building, 4) data exploitation. The recording protocol is couched within the framework of the IPFC program: (i) Text reading; (ii) Guided conversation with a native speaker; (iii) Semi-constrained interaction between two learners. The annotation procedure included sound-aligned orthographic transcription with the Transcriber software, as well as ad hoc annotation conventions. The two most important annotations were respectively dedicated to non-targetlike phonetic-phonological segmental productions (at the word level, but without detailed identification or localization), and to non-targetlike lexico-grammatical sequences, for which a possible targetlike sequence was added in parallel by the native annotator. Thanks to our cooperation with Dr Fontan and Dr Le Coz (Archean Labs, France), a dedicated parser was built to accommodate the annotations and provide overall figures as well as a concordancing data-mining tool. The data was partly explored to investigate different aspects of the speech of the Japanese learners, such as liaison, schwa and fluency on the phonetic-phonological level.

4 . 研究成果

We can distinguish three types of achievements: 1) Achievement of the overall research plan; 2) Quantitative achievements; 3) Scientific achievements.

1) As far as the overall research plan is concerned, we managed to reach our general objective, even though we had to reduce our scope because of technical problems,

human resource management and times constraints.

2) The quantitative dimension of the project is probably the most important achievement, since we managed to gather full data (metadata, text reading, guided conversation with a native and semi-constrained interaction with a peer learner) for 41 students of B1-C1 level across four universities in Tokyo and Fukuoka (essentially 3rd and 4th year university students), out of which 17 were fully recorded a year later (which is sometimes difficult to achieve with 4th year students). The data amounts to 25 hours and 30 minutes (214'278 words) of spontaneous speech plus 58 text-reading files (approximately 3'30 per text, i.e. 3 hours and 40 min). Conversations were orthographically transcribed and aligned with the sound signal using the Transcriber software and manually annotated with ad hoc conventions, resulting in 75'873 annotations, while some of the text-reading files were annotated with the Praat software for liaison and schwa analysis using the IPFC coding system.

3) Regarding the scientific achievements, apart from the database construction and the first exploitations of the corpus used in different papers and presentations, the project managed to offer a first overview of the main learning difficulties that pre-advanced Japanese learners of French still struggle with in spontaneous speech. Looking at the overall corpus, the lexico-grammatical substitution hierarchy (from 1 to 194 according to descending numbers of occurrences) can be presented as such for the top five substitution patterns, pointing to the importance of mastering grammatical determiners in furthering spoken proficiency: 1) le → NULL (180 occ.), 2) la → NULL (150 occ.), 3) le → la (148 occ.), 4) de → NULL (118 occ.), 5) l' → NULL (81 occ.).

Nevertheless, the amount of data and information contained in the database still remains to be fully exploited, and therefore offers itself as a foundation for further research.

5 . 主な発表論文等

[雑誌論文](計 7 件)

- . Detey, S. & Racine, I. (2019). Learning to read liaisons in French as a Foreign Language: a longitudinal study of beginner learners in Japan. *Language, Interaction, Acquisition* 10(1), 97-117. 【Refereed】
- . Fontan, L., Le Coz, M. & Detey, S. (2018). Automatically measuring L2 speech fluency without the need of ASR: a proof-of-concept study with Japanese learners of French. *Proceedings of Interspeech2018* (Hyderabad, India), 2544-2548. 【Refereed】
- . Kondo, M., Fontan, L., Le Coz, M. & Detey, S., Konishi, T. (2018). 英語母語話者と機械による自動評価に基づく 日本語話者の L2 英語の流暢さの評価 [Fluency ratings of Japanese accented English by machine and native English speakers]. *Proceedings of the 32nd meeting of the Phonetic Society of Japan* (Okinawa, Japan), 280-284. 【Refereed】

- . Le Coz, M., Detey, S. & Fontan, L. (2017). Caractérisation automatique du rythme parolier d'apprenants japonophones de FLE. *Special Issue of Revue Japonaise de Didactique du Français (Proceedings of 4th CAP FIPF Congress in Kyoto)*, 504-1-504-8. **【Refereed】**

- . Detey, S. & Racine, I. (2017). Towards a perceptually-assessed corpus of non-native French: the InterPhonology of Contemporary French (IPFC) project illustrated with a longitudinal study of Japanese learners' /b-v/ production. *International Journal of Learner Corpus Research* 3(2), 225-251. **【Refereed】**

- . Detey, S., Fontan, L. & Pellegrini, T. (2016). Traitement de la prononciation en langue étrangère : approches didactiques, méthodes automatiques et enjeux pour l'apprentissage. *Traitement Automatique des Langues* 57(3), 15-39. **【Refereed】**

- . Detey, S., Kawaguchi, Y. & Kondo, N. (2015). La liaison chez les apprenants japonophones avancés de FLE : étude sur corpus de parole lue et influence de l'expérience linguistique. *Bulletin suisse de linguistique appliquée Vals-Asla* 102, 123-145. **【Refereed】**

[学会発表](計 6 件)

- . Detey, S., Fontan, L., Le Coz, M. & Jmel, S. (2018). Application d'algorithmes d'évaluation automatique de la fluence à un corpus longitudinal A1-B1 de parole lue d'apprenants japonais de FLE: perspectives didactiques. International conference *AFLS2018: Focus on French language usage and users*. Toulouse : University of Toulouse-Jean-Jaures (France), 27/06/2018.

- . Detey, S., Le Coz, M., Fontan, L., Barcat, C., Kawaguchi, Y., Akihiro, H., Sugiyama, K. & Kondo, N. (2018). Annotations minimales multi-niveaux d'un corpus de parole spontanée d'apprenants japonais de FLE et traitement automatique : perspectives didactiques. International conference *FLORAL-IPFC2018: contact de langues et (inter)phonologie de corpus*. Paris : MSH (France), 22-27/11/2018.

- . Le Coz, M., Detey, S. & Fontan, L. (2017). Caractérisation automatique du rythme parolier d'apprenants japonophones de FLE. International conference *Congrès CAP-FIPF Ecologie du français & diversité des langues*. Kyoto : University of Kyoto (Japan), 20-24 sept. 2017

- . Detey, S. (2016). *L'impact de l'orthographe sur la prononciation en FLE (illustration avec des apprenants japonais)*. Invited Conference, Dipartimento di Studi Europei, Americani e Interculturali. Roma: University of Roma – La Sapienza (Italy), 12/3/16.

. Lyche, C. & Detey, S. (2016). Enseigner la variation diatopique en FLE dans deux contextes L1 variationnels distincts : le cas de la Norvège et du Japon. International conference *Variation, plurilinguisme et évaluation en FLE - 125 ans d'enseignement du FLE à Genève*. Geneva : University of Geneva (Switzerland), 22-24/6/16.

. Detey, S. (2016). La prononciation des apprenants japonais de FLE : introduction. Table Ronde « La prononciation des apprenants japonais de FLE : caractéristiques et approches. International conference *Congrès de printemps de la SJDF « La parole des apprenants de FLE au Japon : évaluer, corriger, guider »*. Tokyo : Waseda University (Japan), 27-28/05/2016.

〔図書〕(計 件)

〔産業財産権〕
出願状況(計 件)

名称：
発明者：
権利者：
種類：
番号：
出願年：
国内外の別：

取得状況(計 件)

名称：
発明者：
権利者：
種類：
番号：
取得年：
国内外の別：

〔その他〕
ホームページ等

6 . 研究組織

(1)研究分担者

研究分担者氏名：近藤 真理子

ローマ字氏名：Kondo Mariko

所属研究機関名：早稲田大学

部局名：国際学術院

職名：教授

研究者番号(8桁)：00329054

研究分担者氏名：杉山 香織

ローマ字氏名：Sugiyama Kaori

所属研究機関名：西南学院大学

部局名：文学部

職名：准教授
研究者番号（8桁）：00735970

研究分担者氏名：川口 裕司
ローマ字氏名：Kawaguchi Yuji
所属研究機関名：東京外国語大学
部局名：大学院総合国際学研究院
職名：教授
研究者番号（8桁）：20204703

研究分担者氏名：川島 浩一郎
ローマ字氏名：Kawashima Koichiro
所属研究機関名：福岡大学
部局名：人文学部
職名：教授
研究者番号（8桁）：60389335

研究分担者氏名：秋廣 尚恵
ローマ字氏名：Akihiro Hisae
所属研究機関名：東京外国語大学
部局名：大学院総合国際学研究院
職名：准教授
研究者番号（8桁）：60724862

(2)研究協力者

研究協力者氏名：Corentin Barcat, Lionel Fontan, Maxime Le Coz
ローマ字氏名：Corentin Barcat, Lionel Fontan, Maxime Le Coz

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。