

日本人児童英語発話の複雑性及び連語使用の発達 —コーパス言語学と動的システム理論の観点から—

執筆者：小池生夫、鈴木駿吾

研究目的・背景と意義

第二言語習得研究において、発話の複雑性は、正確性や流暢性と同様にスピーキング能力の1つとされている（Suzuki & Kormos, 2020）。しかし、縦断的研究の不足から、学習者の複雑性における発達順序については十分な研究はなされていない（Lambert & Kormos, 2014）。

スピーキングにおける複雑性には、大きく分けて文法的複雑性と語彙的複雑性の2つに分けられる。文法的複雑性は「特定の文法単位内の情報量」と定義される。例えば、「1文に含まれる単語数」という観点から分析が可能である。習熟度が上がるほど、1文に多くの語を含めることができ、一度に言語化できる意味情報の多さを示唆する。語彙的複雑性には、多様さや洗練さなど多くの下位概念が存在するが（Eguchi & Kyle, 2020）、本研究では多様さに焦点を当てる。発話時の多様な語彙使用は、その学習者の語彙知識がより発達していることを示す。これらの文法的・語彙的複雑性という発話能力に対する見方には、「言語が文法的規則に基づいて産出されるものである」という前提がある。しかし、言語使用者は、文法規則だけでなく、決まり文句や定型表現（i.e., 用法基盤）を組み合わせて発話を行うという考え方もある（Ellis & Wulff, 2015）。近年の第二言語習得研究では、「チャンクなどで表現を覚えてから、表現内の文法規則や語彙を習得していくのか」或いは、「文法規則や語彙を習得すると、それらを応用した表現が定着していくのか」といった文法・語彙・用法の発達順序について注目が集まっている。

本研究では、小池生夫氏の博士論文に用いられた年齢の異なる3人の児童の縦断的コーパスに基づき、実際の英語圏での生活における、文法的・語彙的複雑性の発達に加えて、用法基盤の考えに根ざす連語使用の変化も縦断的に描写する。

分析手法

言語指標：（1）文法的複雑性

「T-unit 1つあたりの平均語数」という指標を用いた。本研究では、通常の文ではなく、統語構造に基づく T-unit (Terminal Unit) を文単位とした。T-unit とは、1つの独立節とそれを修飾する従属節全てを含む文単位である（Foster et al., 2000）。

言語指標：（2）語彙的複雑性

先述の通り、本研究では語彙使用の多様さに焦点を当てる。言語指標には、「異なり語の割合」を算出した。指標算出には、TAACO (Crossley et al.,2019) を使用。

言語指標：(3) 連語使用

連語使用の指標に関しては、発話運用能力と関連の高い「2語・3語連語の使用割合」を使用した (Tavakoli & Uchihara, 2020)。上記の指標は、Corpus of Contemporary American English (COCA) の話し言葉データにおける最頻出の 30,000 連語の発話内での使用割合を示す。指標の算出には、TAALES 2.0 (Kyle & Crossley, 2015) を使用。

研究デザイン

発話データは参加者 S, J, N (渡米時年齢はそれぞれ 5、7、10 歳; J, N は男児、S は女児) から採取したものである。発話データは、1972 年 10 月から 1975 年 3 月までの 30 ヶ月のデータを用いた。3 ヶ月毎の区分を各 Stage とし、発話データを 10 Stage に区切り、指標の縦断的变化を描写する。

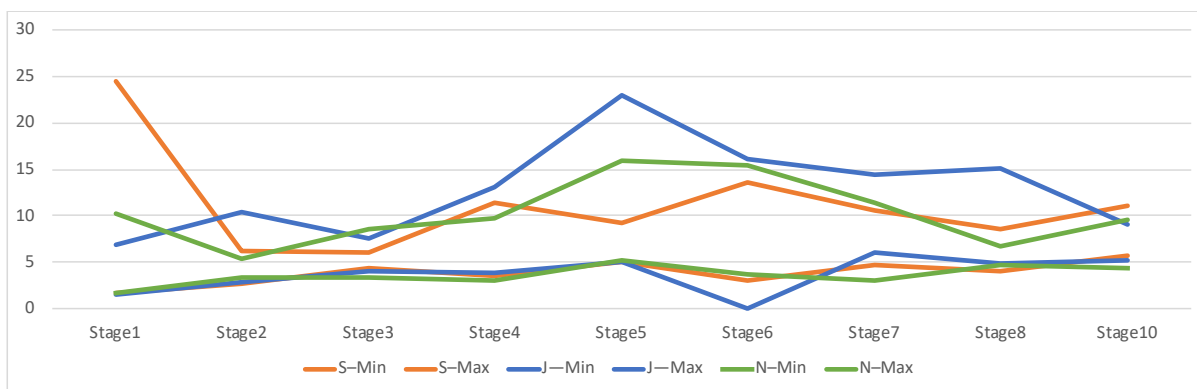
Min-Max plot

本研究では、個人の発話運用の発達を可視化するために、Min-Max plot を用いる (Verspoor, 2011)。Min-Max plot では、各時点における指標の最小値と最大値をプロットする (図 1—4 参照)。動的システム理論に基づき、Min-Max plot では最小値と最大値の幅広さ・ばらつきの時系列変化に着目をする。発達段階初期は、新しい言語知識を取り組み始めている時期のため、言語使用が安定せず、最小値と最大値の幅が比較的広くなるとされている。一方で発達段階終盤には、取り込み始めた言語知識が定着し、言語使用が安定するため、最小値と最大値の差の縮小として現れる。こうした理論的前提に基づき、各参加者の言語指標において、最小値と最大値の拡大・縮小する時期に注目していく。尚、全ての Min-Max plot において、S (渡米時 5 歳; 女児) をオレンジ色、J (渡米時 7 歳; 男児) を青色、N (渡米時 10 歳; 男児) を緑色で描写する。データクリーニングの結果、Stage 9 (1974 年 10—12 月) には 1 つの発話データのみが残り、最小値と最大値が同値となるため、Stage 9 を分析から除外した。

結果・考察

(1) 文法的複雑性—文単位の長さの推移

図 1 : T-unit 1 つあたりの平均語数の推移

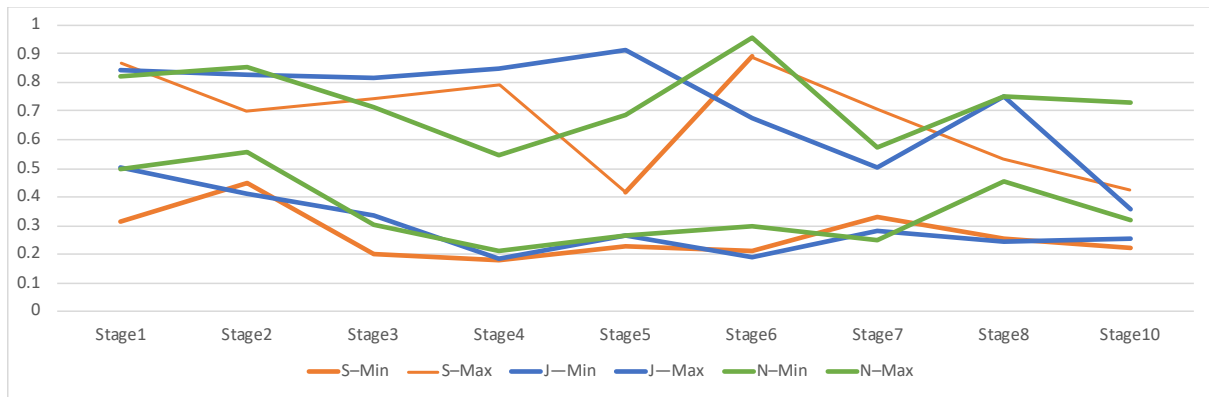


最年少である S は、Stage 2/3・5・8 で最小値・最大値の差に縮小が見られ、Stage 2～3つ程度（6-9ヶ月）で発達が起きていくと考えられる。Jにおいては、Stage 3・7で最小値・最大値の差の縮小が見られるため、1つの発達段階において Stage 4つ程度の期間（12ヶ月）がかかっていたように思われる。最年長である N は、全体を通して比較的最低値・最大値の差が安定しているようだが、Stage 2・8での縮小が見られる。このことから、Nの1つの発達段階には、Stage 6つ分の期間（15ヶ月）がかかっていると考えられる。これらの結果から、学習者の年齢が高くなるほど、新しい言語知識を取り込み始めてから、定着させるためにより時間がかかることが伺える。このことは、年齢が高いほど、より認知的能力が高いということから解釈が可能である。年齢が低い S は、他の J・N に比べると認知的能力が比較的低い状態だと考えられ、新しい言語規則や語彙に関して、一度に注意を向けられる量が少なかったと想定可能であろう。言い換えれば、Sにおいては1つの発達段階において定着させる知識の量が少なかったために、1つの発達段階の期間が比較的短かったと考えられる。一方で、最年長である N は、3人の中で比較的認知的能力が高く、一度に注意を向けられる新たな言語知識の量が多かったために、実際の言語使用としての定着に時間がかかったと考えられる。

これらのデータから、様々な言語特徴に触れることができる英語圏での英語学習の状況では、学習者の認知的能力と、定着可能な言語知識量、その知識の定着にかかる学習時間が密接に関わっていることが伺える。

（2）語彙的複雑性—異なり語の割合の推移

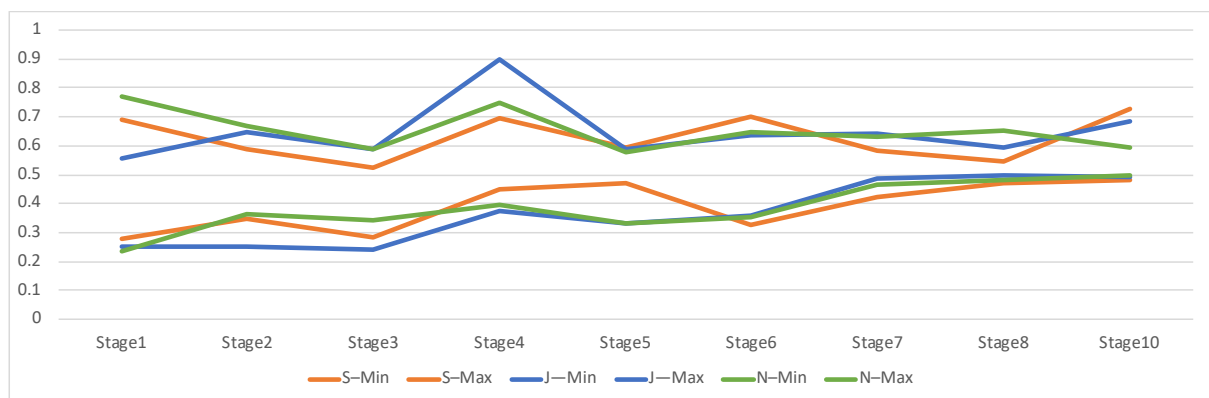
図2：異なり語の使用割合の推移



3人に共通して Stage 1~4 にかけて最小値が減少し、それ以降は最小値が固定化していることが伺える。一方で、最大値の変化には、上昇下降のタイミングの違いはあるものの、時間が経つにつれて下降傾向にあることが伺える。これらのことから、彼らの語彙使用に関しては、ある一定の異なり語の割合に収束していくということが言えよう。つまり、彼らの発話において常に異なり語の割合が増え続けていくというよりも、発話内で一定の同じ語彙を巧みに使いまわすことで円滑な発話やコミュニケーションを実現しているのではないかと推測できる。ただし、異なり語の割合の最大値については、参加者間で上昇下降のタイミングの違いが見られた。語彙知識の発達・増加には、どのような語彙に・どのような頻度で・どのような使われ方で、語彙のインプットを受けているかが重要とされている (Crossley et al., 2013)。参加者の渡米時の年齢を考えると、それぞれ米国在籍時に通学している学校の学年が異なっているはずである。最年長である N は、日常の具体的な話題から、授業科目に関連した抽象的な内容に関する様座な語彙に触れていたことが伺えるし、最年少である S は渡米初期には日常的な会話によるインプットが多かったのかもしれない。語彙学習の発達の仕方に関しては、そうした学習者のインプットを精査した後続研究が必要であることが上記のデータから伺える。

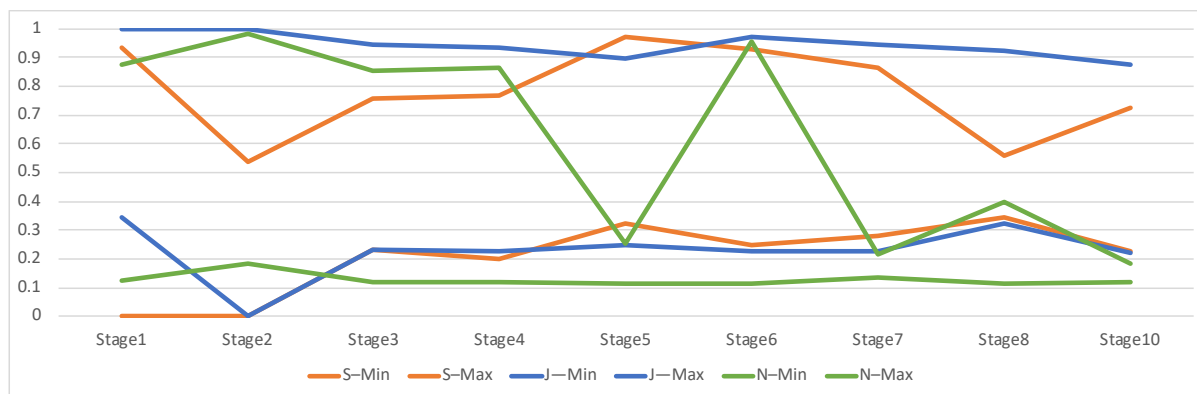
(3) 連語使用—2語・3語連語使用の割合の推移

図3：2語連語の使用割合の推移



2語連語の使用割合については、参加者 S, J, N 全員が、Stage 10 に向かうにつれて、緩やかに最小値が高くなりつつ、最小値・最大値の差が縮小している。このことから、習熟度が発達していくにしたがい、より安定して2語連語を発話内で使用していることが伺える。これは連語使用が発話のスムーズさや流暢性をサポートすること (Tavakoli & Uchihara, 2020) や、習熟度が高い学習者ほどネイティブが用いる連語を使用する (Eguchi & Kyle, 2020) とする先行研究を支持する。

図4：3語連語の使用割合の推移



一方で3語連語の使用に関しては、参加者3人の中で異なる発達の傾向が見られた。最年少のSには、最小値・最大値の差が一定のまま、3語連語の使用の割合がゆるやかに増加していき、Stage 8で最小値・最大値の差が縮小している。一方で、Jは全体を通して、ゆるやかに最小値が向上している。最年長のNは、一定の最小値を保ちながら、最大値はStage4~7で変動し、それ以降で急激に減少し安定している。これらから、2語連語の使用割合とは対照的に、3語連語の使用割合に関しては、Stage後半で比較的安定が見られるも、体系的な発達の傾向は見られなかった。考えられる理由としては、3語連語の使用の定着には、2語連語使用の定着が必要条件なのかもしれない。言い換えれば、定着した2語連語に、学習者の持つ個々の語彙知識を組み合わせることで、3語連語を学習者の語彙知識（心的辞書）に確立させていくのかもしれない (Sprengr, Levelt, & Kempen, 2006)。

結論と総括

(1) 文法的複雑性

学習者の年齢が高くなるほど認知的能力が高くなり、一度に言語体系に取り組む、或いは処理する文法知識・言語情報が増えるため、新しい言語知識が発話として使えるよう定着するまでに長い時間がかかる。

(2) 語彙的複雑性

実際の英語使用では、語彙の多様さを向上させると同時に、一定量同じ語彙を繰り返し使い回していることがわかった。語彙学習のパターンは、実際に使用する語彙使用（読解・聴解も含む）が大きく関わる可能性が示唆された。

（3）連語使用

実際の英語発話では、発話者は文法に基づく算出だけでなく、実際に広く使われている連語を用いて発話をしている。また3語連語使用の発達には、2語連語使用の定着が必要条件である可能性が高い。

限界点と今後の展望

第一に、先行研究では話し言葉は、書き言葉に比べ複雑な言語特徴は発話されにくいとされている。本研究のデータでもそうした話し言葉の性質は例外ではなく、全体を通して、統語的複雑性比較的lowだった。こうした話し言葉の性質が、学習者の統語的複雑性の上限を引き出しきれていない可能性がある。そこで、同様の研究が日記などの書き言葉などのデータでもなされるべきである。第二に、コーパスデータの性質上、タスクパフォーマンス研究のように Stage 間で発話を引き出したタスクやトピックを厳密に統制してはいない。本分析で用いた指標はタスクパフォーマンス研究で使われているものであり、今後はコーパスデータに適した指標の開発が求められる。最後に、本研究では複雑性に関して調査を行なったが、将来的には正確性・流暢性とも組み合わせることで、英語学習者の発達に関してより多面的な知見が得られるだろう。

参考文献

- Crossley, S. A., Kyle, K., & Dascalu, M. (2019). The Tool for the Automatic Analysis of Cohesion 2.0: Integrating semantic similarity and text overlap. *Behavior Research Methods*, *51*(1), 14–27.
- Crossley, S. A., Subtirelu, N., & Salsbury, T. (2013). Frequency effects or context effects in second language word learning. *Studies in Second Language Acquisition*, *35*(4), 727–755.
- Eguchi, M., & Kyle, K. (2020). Continuing to Explore the Multidimensional Nature of Lexical Sophistication: The Case of Oral Proficiency Interviews. *The Modern Language Journal*, *104*(2), 381–400.
- Ellis, N. C., & Wulff, S. (2015). Usage-based approaches to SLA. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: an introduction* (pp. 75–93). New York: Routledge.
- Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring Spoken Language: A Unit for All Reasons. *Applied Linguistics*, *21*(3), 354–375.
- Housen, A., Kuiken, F., & Vedder, I. (2012). *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA*. Amsterdam: John Benjamins Publishing Company.
- Kormos, J. (2006). *Speech production and second language acquisition*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Kyle, K., & Crossley, S. A. (2015). Automatically Assessing Lexical Sophistication: Indices, Tools, Findings, and Application. *TESOL Quarterly*, *49*(4), 757–786.
- Lambert, C., & Kormos, J. (2014). Complexity, accuracy, and fluency in task-based L2 research: Toward more developmentally based measures of second language acquisition. *Applied Linguistics*, *35*(5), 607–614.
- Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, *15*(4), 474–496.
- Sprenger, S. A., Levelt, W. J. M., & Kempen, G. (2006). Lexical access during the production of idiomatic phrases. *Journal of Memory and Language*, *54*(2), 161–184.
- Suzuki, S., & Kormos, J. (2020). Linguistic dimensions of comprehensibility and perceived fluency: An investigation of complexity, accuracy, and fluency in second language argumentative speech. *Studies in Second Language Acquisition*, *42*(1), 143–167.
- Tavakoli, P., & Uchihara, T. (2020). To what extent are multiword sequences associated with oral fluency? *Language Learning*, *70*(2), 506–547.

Verspoor, M. H. L. W. (2011). *Dynamic approach to second language development*. Philadelphia: John Benjamins Publishing Company.