

日本人児童の第2言語としての英語習得
のプロセス研究

2011,12,13 年度
英語教育研究センター
委託研究成果報告

公益財団法人 日本英語検定協会

研究代表者 小池 生夫

共同研究者 投野 由紀夫

蒲原 順子

2014年3月31日

日本人児童の第2言語としての英語習得のプロセス研究

はじめに

「日本人児童の第2言語としての英語習得のプロセス研究」の経過と目的と意義

本研究は過去40年にわたり発展を遂げてきた第2言語習得研究の嚆矢にあたるもので、我が国では文字通り本研究分野の夜明けとなった研究である。

英語の自然環境に一定期間過ごした3人の日本人の子どもがどのようにして英語を自然に習得したかを計画的に記録し、分析をおこない、その特徴を明らかにし、第2言語習得データの分析の背後にどのような言語習得のメカニズムが働いているかを明らかにする研究目的をもっておこなわれたものである。

この研究には、これで完結したものではなく、いまだ研究過程の途中にあるが、いくつかの記録しておくべき内容がある。ひとつは研究の対象になる発話データが2年半の長期にわたること、日本人兄弟妹がそれぞれ10歳、7歳、5歳から2年半の年齢を重ねている在米生活で、当初ゼロの英語コミュニケーション能力が最終的にnative speakerのように英語を駆使できるようになったこと、その記録はnative speakerとの2週間に1回30分のインタビューを中心としてそれぞれ録音され、書写されたものであることである。

また、本研究は、最初の1年の言語記録資料作成と分析は小池生夫の単独の手作業によっておこなわれているが、その後年月をおいてそれを含む2年半全体の資料全体を対象に分析をおこなった。これは投野由紀夫の指導と援助、蒲原順子のコーパス分析に基づくものである。記録、研究のはじまりが1972年10月であり、2013年度末を迎え、42年間の経過がある。第1次研究が終了したのは1979年であり、1980年に米国Georgetown Universityに提出した博士論文は1983年には我が国で出版された英文の著書であり、第2言語習得の研究著書としてはあまりにも早く、後に再販の要望が世間に湧いてきたときには、原版がなく、出版されなかったため、研究者間で「幻の本」といわれたという逸話がある。

小池はその貴重な後1年半の原資料をつづけて分析を試みたが、残念なことに、それに集中できず、そのまま資料分析はなされず、平成10年に小池が明海大学大学院応用言語学研究科を創設する頃、再びこの問題を当時最新のコーパス研究者として帰国した投野由紀夫の勧めと協力があり、ふたたびこのデータを今度は2年半分の全資料をコーパス化するという長い道のりが始まった。やがて、院生のなかから専門的に研究する蒲原順子を研究チームに迎えることができたのは幸運であった。この資料は世界的にも数少ない3人の兄弟妹を対象にした貴重な第2言語発達データであり、日本にはほとんどないものであるため、第2言語習得が実際どのようにおこなわれるのかを知るためには、なくてはならない資料である。しかも、それを知っていながら、研究の基礎である資料にタグをつける作業が進められなかったのは、多忙さと研究資金の不足のためであった。

2011年に公益財団法人日本英語検定協会英語教育研究センターが創設され、その活動の一貫として研究資金を3年つづけて得られたのは、福音であった。研究の基礎作業はふたたび投野と蒲原の協力を得て進みだした。そして、3年計画の最後の年になって、かなり作成したデータをもとに、研究本体が進みだした。

本研究報告は3年分のまとめ、しかも「未完成のまとめ」である。つまり、さらにさまざまなテーマをとりあげて、データ解析をおこない、第2言語習得の過程がどのようなメカニズムであるかを総合的、数量的に分析する状況が生まれている。これから本番というべきか。

本研究報告は、小池が最初の1年の資料を使って形態素、統語構造、意味構造等の習得の順序性が3人の子どもに共通して存在していることを要約した。本研究のイントロダクションである。

また投野が発話の経年変化で発話の統語数、ことなり語の語数などの分析により語彙発達のメカニズムの分

析をおこなっている。また、蒲原は同じく、動詞、助動詞、Be 動詞、名詞などの頻度と傾向とその背景について分析をおこなっている。

1970年代から今日までの長い道のりと研究の手法の発展を利用しての第2言語発達の研究である。とくに過去3年は公益財団法人日本英語検定協会英語教育センターの委託研究に採用され、研究の基礎になるデータのタグ付けがかなり進んで、研究本体が始まったことをここに述べ、深甚なる感謝を申し上げたい。

Part 1 日本人児童の第2言語としての英語の初期の分析 (小池 生夫 担当)

1.1 研究の経緯

1970年の初頭に世界の第2言語としての英語の習得研究ははじまったといえる。これは Harvard University の Roger Brown 博士が心理学的立場から native children が母語を無意識のうちに習得する研究 (1973) が native children の英語の習得に言語学、心理学の面から研究をはじめ、本書が世に出たのである。

小池は、当時米国の Washington, D.C.にある Georgetown University に慶應義塾大学から派遣されて、同大学の School of Languages and Linguistics, Department of Chinese and Japanese の Lecturer として日本語を教える傍ら、Graduate School of Linguistics で博士課程の大学院生として応用言語学を主専攻、理論言語学と英語学を副専攻として単位を取る忙しい日々を送っていた。やがて、書かなければならない博士論文のテーマをどうするべきかを考えていた最中に本書にめぐりあったのである。

この研究対象は、native children であるが、私の子どもたち3人は日本人で、米国に来て、近所の公立小学校に毎日通い、土曜日は米国大使館で開かれる日本人補習学校に通う生活をしてきた。家の中では兄弟妹は日本語で両親と日本語で話したから、ほとんど完全な母語環境であるが、一旦外にでると学校の生活や近所の子どもたちと遊ぶ場合は完全に英語環境であった。このような生活を2年半つづき、日本に帰国した。この間、どのようにして英語を自然に習得するに至るかについて私は研究対象として前人未踏の世界を選び、日本人の3人の子どもで、渡米時10歳、7歳、5歳でお互いにはほぼ3年近く年齢差がある子どもで、上2人は男児、小さい子は女児であった。子どもの英語を自然環境で取り扱って、英語を教える世話をすることはしなかった。そのままにした結果どのようなことがおきるのかを研究することにしたのである。それは1972年9月から1975年3月までの2年6ヶ月の期間である。

3人は Nobi (N), Jun (J), Sachiko (S) と研究上命名した。前述したように、その後第2語習得研究は世界的に大発展したが、当時はほとんどだれも関心を示さず、たまに日本でも帰国子女について、英語やその他の住んでいた国、地域の言語が流暢に使えて、教室では英語の教師たちが自分よりはるかに発音が native のようで、流暢にできる子どもの取り扱いに困るといような噂が聞こえる程度であった。なにが、どのようになぜ、そうなるのかについて本来積極的に研究方法を開発すべき言語学、言語教育研究者は研究の方法も判らず、傍観し、話のタネにする程度にすぎなかった。これが当時の世界のこの言語習得研究の状況で、たとえば「夜明け前」であった。

私は、当時40歳、英語の教師の経験は18年あり、このテーマについて、渡米前から大変興味があった。まるで「未開の扉を開ける」ような知的興味があったが、Georgetown University に行ってはじめて、応用言語学研究の世界のメッカである環境から、いろいろの研究上の示唆を受けることができた。具体的に取り組み、できれば博士論文としてみたいという気持ちが湧いたのである。その際に具体的に研究の目的、方法を先の A First Language (1973) や Hakuta (1974) などを読み、参考にした。

1.2 具体的分析の方法

1人ずつの子どもを30分ずつ、**native adult**の**interview**を受けてお話しをする。テーマはさまざまであるが、おもに児童用の物語を絵本を見ながら、説明する。これは日本語が多い。できるだけ多くの文章が英語で語られるように仕向ける。しかし、最初は子どもはほとんど話せない。時間は3ヶ月ごとに区切り、2週間に1回の**interview**は30分とする。その英語の会話を録音し、聞いて書写する。その際に間違いを発見し、その正用法を書き加える。そして、間違いを明確に記録する。そのような単純作業を継続した。そののちに、エラーの分析をおこなった。ひと月に2回、それを3ヶ月で6回分のデータを比較検討すると英語の習得の変化現象が明瞭になる。一ヶ月に2回の資料では変化現象として指摘できない。

1.3 中間言語、エラーの分析と習得の過程

子どもたちの英語は、単語一語、または語の連結からはじまる。音の連結は、意味、構造、機能を持つが、常に完全ではない。それは中間言語 (**inter-language**) である。中間言語は **native language** のような完全な言語ではなく、形態に不足または間違いあるいは過大な要素が加わったものである。例えば、**two childs** というのは、正しくは **two children** である。この間違いは過大な連想によって **child** の複数 **dog, cat, box** の複数と同じ複数化の規則に従うと発想してしまうことによる間違いである。

そのような過程を経て、子どもたちは **interlanguage** の過程から **native language** の過程に英語を習得するに至る。理想的に言えば、100%正確に流暢に聞き、話し、読み、書くレベルになって、はじめて英語を習得し、**native speaker** と同じ段階に到達する。しかし、そこまで行く間に10例のうち80%正確に流暢に発話できた時をもって、その形態素、語彙、文法形態が習得されたと仮定することにする。その決め方は研究者によって異なる。**Brown** は90%といい、**Slobin** は一例でも無意識に正しく発話されれば、その時点でその語彙は習得の域に達したと仮定する。小池はその中間的な立場をとるが、録音し、書写した言語資料全体を通してのバランスで判断した。

1.4 形態素、語彙、文構造、意味構造の習得とその順序性

発話資料を分析した結果、発見したことは、第2言語の形態素、統語構造の習得のプロセスではどの子どもも共通の同じ順序で習得がなされるという仮説である。これは、形態素群、語彙、文構造において3人の子どもはほぼ同じ順序で英語の習得を無意識に行っていくという仮定である。これは、**Krashen (1978)** のいう **Input hypothesis** から支持を受けたものである。

1.5 第2言語としての英語の習得とは

3人の子どもの言語の習得とは、子どもが無意識に **native** と会話している状態で記録された文や形態素の中でその形態素や文構造が10例引き続いて現れて80%以上の正用率を示すのに達した時点で習得 (**acquire**) されたと仮定する。10例以上にならない場合でもそれが、8例あれば、「習得に近い」状態と仮定する。正用率は研究者によって異なる。90%の正用率を示した場合であるとか、一例でも正しい用例があれば、習得の域に達したとする研究例もある。

1.6 記録の方法と規模

会話は音声記録したもので、筆写しなければ分析できない。それも筆写は日本人には無理である。実際に

発音、イントネーションなどが **native** に近いレベルになっているか、表現が正しいか、文法的間違いがあるかなど、特に発音が判断できるのは **native speaker** でなければならない。私はその点もとても信頼できる米国婦人にめぐりあった。子どもの友達で親しくなった家族ごとの交際をしてくれた家庭の主婦であるが、すばらしい能力をもっていた。そのひとが会話の相手から書写まで間違いを的確に判断してくれたおかげで良質の書写ができたのである。

書写は記録した音声テープを 2 度ずつ確認し、生の発音に基づいて **spelling** を造ってくれた。(たとえば、**This toy is good.** は *Dis toy 'z good.) 音声録音されたテープは 2 年と 6 ヶ月に 60 分用テープは 130 本に達した。書写は帰国してからも米国からの留学生 **native** に頼み、その作業は 4 年の歳月を費やした。全部で 790 頁のタイプした資料と手で書いた資料である。

録音テープは発音チェックが 4 本、大人と子どもの会話が 32 本、子どもと子どもが 63 本、ストーリーを語るテープが 17 本、テストとコミュニケーション行動は密接に結び付くのである。

1.7 3 人の子どもの形態素群の習得の順序性

Spearman の系列相関値計算によると 3 人の子どもは概ねすべての形態素の習得順序に相関が強くまたは弱くある。それをすべて紹介する余裕はないが、事例として、代名詞 **I, you(主格), this(主格), we, my, that, he, she, this(指示詞), you(目的), it(主格), me** の 12 代名詞での習得の順序では、**I** が最も早く習得される。5 歳児と 7 歳児に 0.61 の準強の習得の順序性がある。つまり、よく似た順序で習得がなされている。5 歳と 7 歳では 0.78 の習得の順序性が強く存在する。7 歳児と 10 歳児では 0.84 の強い相関がある。

助動詞群ではどうか。can(ability, possibility), don't, can't (possibility), could(possibility), be gonna, will (II) の 6 助動詞では 5 歳と 7 歳に 0.60、5 歳と 10 歳に 0.49、7 歳と 10 歳に 0.77 の相関性がある。5 歳と 10 歳児は弱い相関がみられるが、後は相関性がみとめられる。

前置詞 in(locative), on(locative), to(goal), at (locative) には、in が最も早く、習得されるのであるが、5 歳児と 7 歳児では 0.5 で弱い相関があり、5 歳児と 10 歳児では 0.8 で強い相関がある。7 歳児と 10 歳児では 0.3 で相関は弱い。

ところが、**Be** 動詞系では、いずれも相関が高い。is, am, are (you, sing), was(he), are (they) では 3 人とも 0.96 から 1.00 の極めて高い数値がでている。

1.8 統語構造の習得の順序と相関

今まで形態素の相関性をみてきたが、大体は 3 人の子どもの間に相関が見られた。そこで、彼等が文表現を表現できるようになったときに、その習得順序の相関性がどうなるかは興味があるテーマである。なぜなら、だれもこの問題に手をつけられないからである。

表1 3人の子どもの統語構造の習得順序

番号	文構造	5歳児	7歳児	10歳児
1	NP-V t -NP	3	(3)	(3)
2	NP-Vi	5	4	3.5
3	NP-Be-NP	3	1	3.5
4	NP-Vt-NP	1	(2)	1
5	NP-Vt-NP	6.5	5	5
6	NP-Vt-Ad(L)	6.5	6	7
7	NP-Be-Ad(L)	8	7.5	7
8	NP-Be-Adj	3	7.5	7
9	NP-Vi-Ad(M)	10	9.5	9
10	NP-Vt-NP-NP	9	9.5	10

注 NP 名詞 Vt 他動詞 Vi 自動詞 Be Be 動詞 Ad(L) 場所を示す副詞 Adj 形容詞
Ad(M) 様相を示す副詞

下に記述したように、Spearman の習得順序系列相関値計算方式による計算では以下のように数値がでる。
(0.7 以上は相関が非常に高いという意味) これは 3 人とも統語構造が使えるようになる順序はほとんど同じ
であるという数値である。これから推量するのに、3 人は同じ順序で無意識にこれら統語構造が使えるように
なることである。

5歳児-7歳児 0.82 5歳児-10歳児 0.85 7歳児-10歳児 0.94

上記の習得順序相関率をみると、統語構造でも 3 人の子どもの相関率は非常に高いといえる。

参考文献

- Brown, Roger W. 1973. A first language: The early stages. Cambridge, Mass., Harvard University Press.
- Dulay, Heidi c. and Marina. k. Burt, 1972. Goofing: An indicator of children's second languages learning strategies, Language Learning 22:2. 235-252.
- Hakuta, Kenji. 1974. A preliminary reports on the development of grammatical morphemes in a Japanese girl learning English as a second language. Working Papers on Bilingualism 3. 1-17.
- Hakuta, Kenji. 1975. Becoming bilingual at age, five: The story of Uguisu. Unpublished senior honors thesis, Harvard University.
- Koike, Ikuo. 1983. Acquisition of grammatical structures and relevant verbal strategies in a second language. Tokyo. Taishukan Publishers
- Krashen, Stephen B. 1977. The monitor model for adult second language performance. In: Viewpoints on English as a second language, ed. By M.K. Burt, H.c. Dulay and M. Finocchiaro. I 52-61, New York, Regents.
- Makino, T. 1981. Acquisition order of English morphemes by Japanese adolescents. Tokyo. Shimozakishorin Press.
- Slobin, Dan Issac and Welsh, c. 1973. Elicited imitation as a research tool in developmental psycholinguistics. In: Studies in child language development, ed. By c, Ferguson and D. Slobin. New York, Holt.

Part 2 小池コーパスの3人の子ども別サブコーパスを用いた発話語彙・統語分析
 (投野 由紀夫 担当)

2.1.小池コーパスの話者別分析

小池コーパスは、3名の日本人英語学習者（児童）による、発話データからなるコーパスである。本コーパスの特徴として、長期縦断的（longitudinal）データであることがあげられる。通常、第二言語習得研究には、その利便性から横断的（cross-sectional）アプローチによって得られたデータを使用することが多いが、横断的アプローチで得られるデータは、いわゆる自然界においての言語習得でなく、異なる学習者の結果を比較しているという問題が残る。本コーパスのような縦断的データを使用した研究は、横断的データによる研究と相互補完的關係にあり重要である。

小池コーパスの基礎統計量を表1、表2に示す。

表1：小池コーパス統計

3人の子供データ	総語数	参加収録数	収録期間の年齢
N	16352 words	88回	11歳2ヶ月～14歳5ヶ月
J	25903 words	104回	7歳8ヶ月～10歳3ヶ月
S	33110 words	104回	5歳4ヶ月～7歳8ヶ月

表2 小池コーパス サブコーパス統計 (単位：総語数)

月数	3	6	9	12	15	18	21	24	28	31	34
N	618	785	1790	3472	2994	1914	2270	720	1301	65	423
J	554	625	1520	3437	6754	3611	2687	1016	3666	838	1195
S	364	1470	4659	6492	5907	3767	2823	1858	3212	1244	1314
総収録数	9	18	25	23	17	15	9	6	5	2	4

収録のスペンは各3ヶ月（25-28ヶ月のみ4ヶ月）となっているが、収録時期、タスク（トピック）などの条件が異なり、また全員が毎回の収録に参加しているわけではない。

表3、は各学習者の参加した収録回数と、1回あたりの発話数（総語数）である。

表3 小池コーパス 学習者の収録参加回数 (単位：回)

月数	3	6	9	12	15	18	21	24	28	31	34
N	8	11	17	14	9	10	8	5	4	1	1
J	8	12	18	20	13	12	7	5	5	1	3
S	8	17	15	18	13	10	8	5	5	2	3
収録回数	9	18	25	23	17	15	9	6	5	2	4

JとSと比べて、Nのみ収録へ参加した回数は若干少ない。最後の一年（24ヶ月～36ヶ月）は収録の回数自体が減少しているが、学習者の言語発達のため一回の収録で得られるデータが増加しているため、発話数自体が極端に減少したわけではない。

2.2. 語彙発達について

表 4、図 1 は、収録 1 回あたりの学習者の発話量をまとめたものである。

表 4 収録 1 回あたりの総語数

月数	3	6	9	12	15	18	21	24	28	31	34
N	77.3	71.4	105.3	248.0	332.7	191.4	283.8	144.0	325.3	65.0	423.0
J	69.3	52.1	84.4	171.9	519.5	300.9	383.9	203.2	733.2	838.0	398.3
S	45.5	86.5	310.6	360.7	454.4	376.7	352.9	371.6	642.4	622.0	438.0

3名のいずれとも、初期から発言量は時間経過とともに漸増傾向にあるが、爆発的に量が増える地点が存在する。N、Jは10-12ヶ月の時点で、Sは7-9ヶ月の時点でそれぞれ、収録一回あたりの発話量が前月比で2倍以上になるなど、発話の総語数が爆発的に伸びている。収録の時間やタスクなどの条件が均一でないため一概に結論付けることはできないが、12ヶ月以降、発話量が極端に減少することはほぼない（Nの29-31ヶ月目のみが例外である）。

個別に観察すると、Nの発話量の増加はJ、Sと比べ少なく、この傾向は12ヶ月以降でも変わっていない。対して最初の1～3ヶ月間の記録では、最も多く発話したのはNであった。これには、渡航時点で11歳2ヶ月というNの年齢が関連していると推測される（Jは7歳8ヶ月、Sは5歳4ヶ月）。即ち、渡航直後は最も認知的に発達している年長として対話に参加し、英語の学習が進むにつれ、他2人が認知的にも成長し、遜色のない発話を行うようになったのであろう。特にSの伸びは爆発的といってよく、6ヶ月時点まででは単語の羅列でしかほぼ対話に参加できていなかったのが、7～9ヶ月の期間には、英文として必須要素である主語・動詞を備えた文を表現できるようになっている。

Token per Recording

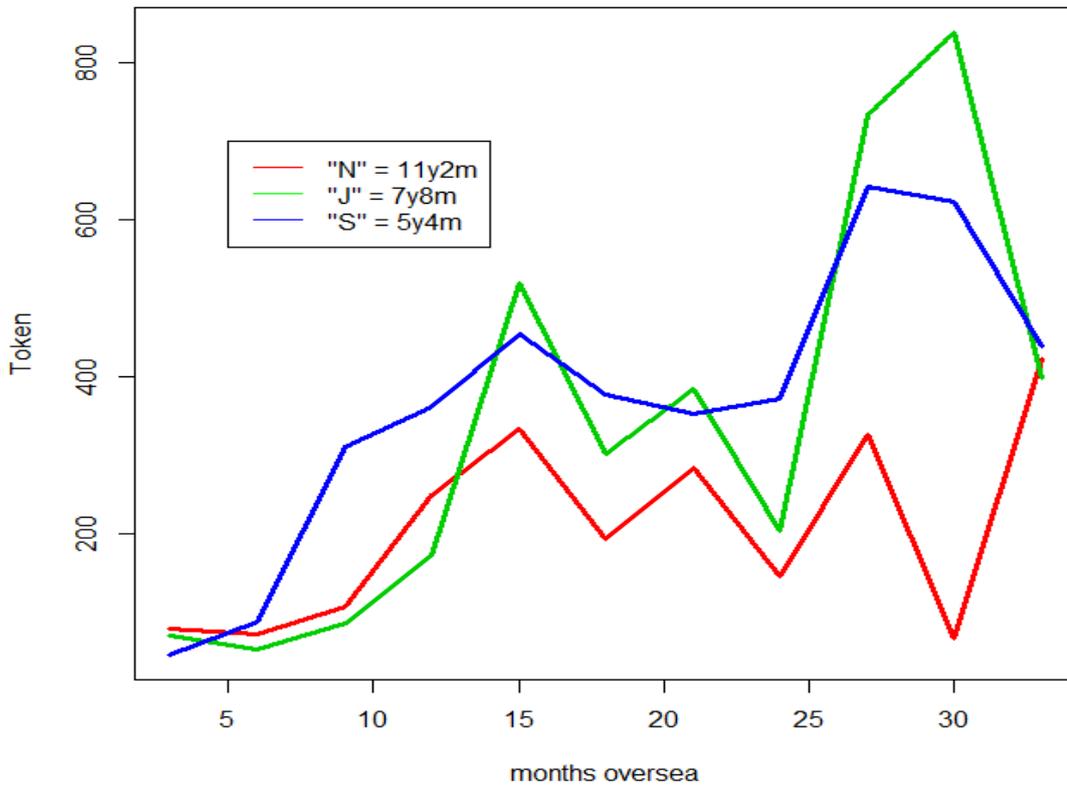


図 1 各学習者の収録 1 回あたりの総語数推移

続いて語彙の豊かさを知るための指標として、総語数、異語数、Type-Token Ratio (異語と総語数の比)、ギロー指数 (異語数を総語数の平方根で割ったもの) を各学習者についてまとめたものが、表 5、表 6、表 7 である。Type-Token Ratio は、テキストの語彙密度を調査するのに用いられる指標で、テキストの総語数 (token) を、異語数 (type) で除算することで求められる。最もシンプルな指標として広く使用される Type-Token Ratio であるが、テキストの長さにより数値が大きく影響を受けてしまう欠点知られており、これを補う目的でしばしば併用されるのが、ギロー指数 (Guiraud Index) である。これは総語数を異語数の平方根で除算することで求められる値で、分母の値を抑えることで、極端に短い、または長いテキストの語彙密度をより正確に表すことで知られる。

表 5 話者 N の基礎語彙統計量

月数	3	6	9	12	15	18	21	24	28	31	34
総語数	618	785	1790	3472	2994	1914	2270	720	1301	65	423
平均総語数/回	77.3	71.4	105.3	248.0	332.7	191.4	283.8	144.0	325.3	65.0	423.0
異語数	352	308	642	655	589	428	536	234	298	51	150
TTR	0.57	0.39	0.36	0.19	0.20	0.22	0.24	0.33	0.23	0.78	0.35
Guiraud Index	14.16	10.99	15.17	11.12	10.76	9.78	11.25	8.72	8.26	6.33	7.29

表 6 話者 J の基礎語彙統計量

月数	3	6	9	12	15	18	21	24	28	31	34
総語数	554	625	1520	3437	6754	3611	2687	1016	3666	838	1195
平均総語数/回	69.3	52.1	84.4	171.9	519.5	300.9	383.9	203.2	733.2	838.0	398.3
異語数	345	230	526	644	981	634	581	333	563	235	333
TTR	0.62	0.37	0.35	0.19	0.15	0.18	0.22	0.33	0.15	0.28	0.28
Guiraud Index	14.66	9.20	13.49	10.98	11.94	10.55	11.21	10.45	9.30	8.12	9.63

表 7 話者 S の基礎語彙統計量

月数	3	6	9	12	15	18	21	24	28	31	34
総語数	364	1470	4659	6492	5907	3767	2823	1858	3212	1244	1314
平均総語数/回	45.5	86.5	310.6	360.7	454.4	376.7	352.9	371.6	642.4	622.0	438.0
異語数	213	369	785	814	857	638	640	450	516	303	364
TTR	0.59	0.25	0.17	0.13	0.15	0.17	0.23	0.24	0.16	0.24	0.28
Guiraud Index	11.16	9.62	11.50	10.10	11.15	10.39	12.05	10.44	9.10	8.59	10.04

最初期の 1-3 ヶ月間の異語数が多く、語彙密度も高いのは、発話中に日本語でのものが多く、異語数にも多量の日本語の単語が含まれていることに起因する。この時期に特徴的な点として、日本語でカタカナ語として使用される単語、数例を挙げると **one, blue** といったような単語を、「ワン、ブルー」といった日本語発音ながら、なるべく多く使おうとする努力が見られた。また、1-3 ヶ月の間、英語の発音から過度に逸脱した場合、日本語ローマ字表記によってアルファベット記述されていることを付記したい (one を wan、blue を bulu: のように記述している)。

語彙密度に関しては、指標が発話の量に大きく影響を受けるため数値が低下しているが、これはむしろ、堅調な語彙発達を裏付けるものとみなすことが可能だろう。実際、その点は異語数が増加しているところからも確認できる。語彙修得の習得過程では、量的に順調な成長のある時期と、語彙的・文法的複雑さを習得する質的な変化のある時期が両方観察される。詳細は後に譲るが、10~12 ヶ月目では、過去形の使用、疑問文 (疑問詞含む)、否定文など様々な操作を必要とする言語特徴が観察され、文法上の誤りも散見されるものの、完全な文の使用が増える時期と重なり、日本語の使用もほぼ見られなくなるのが特徴的だった。

表 8 各話者の異なり語の累積頻度

月数	3	6	9	12	15	18	21	24	28	31	34
N	352	607	1083	1455	1695	1837	2015	2047	2128	2133	2152
J	345	537	919	1325	1857	2076	2254	2306	2459	2490	2553
S	213	553	1088	1499	1846	2028	2200	2285	2411	2449	2506

表 9 各話者の総語数の累積頻度

月数	3	6	9	12	15	18	21	24	28	31	34
N	618	1403	3193	6665	9659	11573	13843	14563	15864	15929	16352
J	554	1179	2699	6136	12890	16501	19188	20204	23870	24708	25903
S	364	1834	6493	12985	18892	22659	25482	27340	30552	31796	33110

表 10 各話者の type-token ratio

月数	3	6	9	12	15	18	21	24	28	31	34
N	0.570	0.433	0.339	0.218	0.175	0.159	0.146	0.141	0.134	0.134	0.132
J	0.623	0.455	0.340	0.216	0.144	0.126	0.117	0.114	0.103	0.101	0.099
S	0.585	0.302	0.168	0.115	0.098	0.090	0.086	0.084	0.079	0.077	0.076

表 8、表 9、表 10 は、語彙的な成長を観察するため、積み上げ方式で総語数・異語数を学習者別にまとめたものである。表 5、表 6、表 7 との違いは、表 5～7 が各期間の異語数を個別に集計していたのに対し、表 8～10 は「積み上げ方式」、つまり前の期間の総語数・異語数を加えて集計した点である。これらから読み取れる傾向としてもやはり、総語数・異語数ともに、最初は $N > J > S$ であったものが、時間の経過とともに逆転するのが見て取れる。最終的な 34 ヶ月目の到達時も、J と S を比較すると、S の方が発言は多いものの、異語数的に J との差はほとんどないことが読み取れる。

2.3. 単語の使用について

3 名の具体的な単語の使用をまとめたのが、表 11、表 12、表 13 である。1～9 ヶ月時点までの発話は、N、J、S3 名ともインタビューアーの問いかけに対する YES, NO の受け答え、及び日本語の使用が中心であり、厳密な意味での連語使用は僅少であったことがわかる。英単語を用いている場合でも、kam(come)、buk (book)、blak (black)、gu bai (good bye) などの表記がされ、日本語のカタカナ語として区別されている。

変化が見られるのは、発話の総量が増加した時点と同時であるのが興味深い。N、J は 10～12 ヶ月の時点、S は 7～9 ヶ月の時点がそれにあたるが、and、the、代名詞などの機能語が増加しているのが観察される。特に and の使用が顕著に多いが、これは本来の等位接続詞の用法以外に、つなぎ言葉として文頭で用いられていることが多いためでもある。

また、16 ヶ月以降のタスクには Story Telling Task などの活動が取り入れられており、内容語で上位に来ているもの (little, Gulliver, wolf など) は、タスクによる影響とみなすことができる。

表 11 N の単語使用上位 10 位

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	39	no	58	i	65	i	174	and	332	and	167	the
2	20	yes	39	no	54	no	136	the	209	the	142	and
3	11	kam	39	yes	51	and	111	is	116	is	53	to
4	11	tip	22	do	49	the	97	i	85	he	45	this
5	9	do	21	you	37	to	80	's	72	's	43	's
6	7	gud	15	n't	29	you	67	this	59	this	37	is
7	7	hi	15	to	27	's	56	no	53	on	34	wolf
8	7	ho	12	know	26	my	55	to	51	there	31	house
9	6	k	12	like	25	is	51	n't	49	i	26	so
10	6	to	11	and	22	do	47	he	38	a	25	he
月数	21		24		28		31		34			
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖		
1	162	the	35	the	68	i	4	this	49	and		
2	105	and	27	i	55	the	3	that	31	the		
3	73	to	25	and	38	you	2	and	24	he		
4	50	that	23	to	37	to	2	but	18	then		
5	49	i	22	's	35	's	2	does	15	to		
6	45	a	22	n't	35	it	2	go	12	said		
7	43	's	18	a	30	n't	2	is	11	pinochio		
8	37	gulliver	18	do	27	that	2	n't	9	i		
9	34	so	18	is	23	we	2	one	7	a		
10	30	n't	15	you	21	have	2	some	7	legs		

表 12 J の単語使用上位 10 位

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	39	no	33	yes	68	and	244	and	473	and	213	the
2	15	yes	23	no	56	is	174	the	353	the	155	and
3	7	:	21	is	51	yes	110	you	192	's	138	's
4	7	to	20	and	49	no	75	i	176	this	88	there
5	6	do	19	i	38	i	74	's	161	is	87	a
6	6	kam	18	you	34	the	72	is	155	a	80	you
7	4	and	13	english	32	my	51	n't	145	he	75	i
8	4	buk	13	have	32	this	48	little	104	to	74	is
9	4	n't	12	do	31	's	45	do	102	you	68	it
10	4	onnanoko	9	one	27	n't	45	yes	92	it	68	n't
月数	21		24		28		31		34			
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖		
1	178	the	54	the	180	the	40	the	57	you		
2	121	and	36	n't	118	and	28	this	45	to		
3	79	i	33	i	114	i	24	it	32	and		
4	72	to	32	you	108	's	23	's	32	the		
5	64	's	28	it	90	it	23	they	30	he		
6	55	a	26	's	81	you	20	and	29	a		
7	55	this	25	do	79	they	19	here	27	's		
8	50	you	25	to	67	she	19	is	19	okay		
9	47	it	17	and	67	to	18	have	18	it		
10	41	he	16	is	58	n't	18	to	17	like		

表 13 S の単語使用上位 10 位

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	32	no	48	this	190	you	307	and	415	and	167	the
2	18	yes	46	you	181	i	194	i	412	the	164	and
3	5	dog	41	i	169	this	180	's	196	's	109	's
4	5	du	39	no	160	and	178	it	168	a	107	i
5	5	pardon	33	is	134	do	177	we	142	is	97	you
6	4	do	32	and	130	it	175	the	118	to	86	it
7	4	faib	29	k	115	n't	171	you	109	there	83	to
8	4	gu	29	o	98	to	163	that	104	i	71	a
9	4	kam	27	yes	93	is	151	this	92	it	69	that
10	4	n	24	do	89	's	127	to	92	you	64	she
月数	21		24		28		31		34			
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖		
1	128	the	61	and	145	the	43	it	56	the		
2	111	and	57	the	108	and	40	the	51	and		
3	105	i	54	you	106	i	40	they	46	i		
4	63	's	52	a	92	to	39	and	45	he		
5	62	a	51	's	71	you	39	i	45	to		
6	61	to	50	i	66	's	35	you	30	you		
7	55	that	43	to	66	she	31	's	19	go		
8	47	it	41	is	63	it	31	this	18	like		
9	47	she	39	n't	48	is	24	is	18	was		
10	42	in	37	it	45	that	21	that	17	a		

2.4. 連語について

本項では、データ分析にシャノン・ヴィーヴァー (1964) によって考えられた N-gram モデルを用いる。N-gram モデルとは、文字列の中で任意の N 個の文字や単語が組み合わされる確率を求めるモデルであり、機械的に言語データを処理することで、人間にとって探し出すことが困難な特徴を発見できる方法として用いられている。

学習初期の 3 ヶ月付近こそ、yes や no、ok といった単純な応答を含む連鎖が高頻度だが、時間の経つにつれ、don't や have to などの機能語や、there 構文、接続詞や代名詞を含む連鎖が上位にくるように変化している。中でも特に and の使用が顕著だが、これは等位接続詞の用法というより、話し言葉というモードゆえのつなぎ言葉としての、文頭の使用が目立つ。また、12 ヶ月以降、タスクも Story Telling Task といった複雑なタスクが混ぜられており、時折上位に現れる内容語 (Gulliver や sambo など) は、それらのタスクにて使用した単語であることが見て取れる。

頻度的に上位にこそ来ていないが、過去形の使用については、3 名とも 12 ヶ月の時点で初めて使用が見られた。特記すべき点として、初出の時点で規則動詞・不規則動詞の過去形が同時に使用されている (made, broke などの不規則活用形と、missed, borrowed などの規則活用形が観察される)。また、誤用も散見されたものの、did not (あるいは didn't) 元型の正しい形での使用がみられた。また、疑問文も少数ながら最年長の J による使用が観察された。12 ヶ月以降も、過去形は 12 ヶ月時点と大差なく使用されている。また現在完了形も 12 ヶ月の時点から使用が見られるが、過去形と比較すると全データで中 12 件と、過去形と比較して少数に留まった。頻度的に見て取れるのはむしろ現在分詞だが、これには話し言葉というモードゆえの多さもあるだろう。

表 14 N の bigram 使用上位 10 位

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	16	no no	13	i do	13	do you	33	do n't	51	and he	37	and the
2	10	yes no	13	yes yes	10	you like	20	this is	44	and the	28	the wolf
3	9	no yes	12	do n't	9	that 's	19	little black	32	on the	20	this is
4	5	hi tip	10	n't know	9	this is	17	black sambo	25	there is	18	house and
5	5	tip hi	9	do you	8	go to	17	pulling the	19	and there	14	is the
6	4	kam hia	9	yes no	8	no i	16	gon na	14	and white	14	went to
7	3	gud bai	7	go to	6	do n't	15	and he	14	in the	12	and he
8	3	nain ten	7	no yes	6	i am	14	i do	14	there 's	9	's house
9	3	wan tu	6	i go	6	i do	14	is pulling	14	this is	9	to the
10	3	za do	6	no no	6	in the	14	tiger said	13	boy is	8	wolf came
月数	21		24		28		31		34			
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖		
1	21	the gulliver	11	do n't	14	have to	2	does n't	15	and then		
2	19	and the	7	it 's	12	do n't	2	one is	8	and he		
3	16	have to	6	's a	12	i 'm	2	this one	8	and the		
4	16	in the	6	and there	11	it 's	1	'm teddy	5	then the		
5	15	the other	6	there 's	11	n't know	1	's your	4	him and		
6	12	do n't	5	they are	9	on the	1	about trudy	4	Pinocchio said		
7	12	to the	5	this is	8	in the	1	and but	4	the Pinocchio		
8	11	on the	4	go to	8	who 's	1	and the	4	then he		
9	10	other island	4	i do	7	that 's	1	back this	3	he put		
10	10	so the	4	is a	6	'm catching	1	better you	3	he went		

表 15 J の bigram 使用上位 10 位

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	18	no no	7	do n't	19	this is	40	and the	67	and he	41	there 's
2	4	and white	7	i do	18	do n't	40	little black	57	and the	36	do n't
3	4	do n't	7	n't know	14	i do	33	black sambo	53	this is	27	and the
4	4	no yes	6	can you	12	n't know	28	do n't	43	there 's	25	it 's
5	4	yes yes	6	no yes	11	that 's	20	that 's	40	and there	20	's a
6	3	i do	6	speak english	10	yes yes	19	in the	38	in the	20	this is
7	3	n't know	6	yes yes	7	name is	18	i 'll	31	's a	18	have to
8	3	to iuno	5	do you	7	no i	17	the tiger	31	and this	16	over here
9	2	bag map	5	i have	6	i like	15	n't know	29	he 's	16	that 's
10	2	balu:n tennis	5	is coming	6	this one	14	and little	27	and then	14	over there
月数	21		24		28		31		34			
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖		
1	31	and the	16	do n't	23	have to	7	have to	14	would you		
2	17	it 's	11	i do	22	let 's	7	it 's	14	you like		
3	15	do n't	8	n't know	21	i 'll	7	this is	11	like to		
4	15	have to	7	have to	21	she 's	6	let 's	11	to say		
5	15	in the	6	's your	20	it 's	5	's the	10	and he		
6	13	to the	5	did n't	19	do n't	5	do n't	9	say something		
7	12	the barber	5	it 's	16	and then	4	going to	8	i 'm		
8	11	and then	5	n't have	16	gon na	4	gon na	8	it 's		
9	11	i know	5	the fire	16	just pretend	4	n't have	8	thank you		
10	10	on the	5	you have	16	this is	4	the kitchen	8	to go		

表 16 S の bigram 使用上位 10 位

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	15	no no	29	o k	65	do n't	50	do n't	82	there 's	26	that 's
2	6	no yes	18	do n't	57	this is	49	it 's	73	and the	24	it 's
3	4	yes no	15	this is	49	o k	41	gon na	59	and then	23	do n't
4	3	no pardon	12	no no	40	have to	41	we 're	57	on the	23	then the
5	2	ap do	11	ma ma	39	ca n't	38	like this	45	's a	22	i think
6	2	berj gud	8	for you	27	like this	35	have to	45	and there	19	she said
7	2	boj grl	8	yes yes	26	it 's	34	and then	26	in the	18	in the
8	2	dadi gu	7	round and	25	i do	29	that 's	25	's the	15	and the
9	2	du ju	7	that 's	24	do it	26	ca n't	22	this is	15	the little
10	2	du zat	6	is for	21	you ca	25	and the	20	and he	14	then she
月数	21		24		28		31		34			
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖		
1	17	in the	18	do n't	27	i 'll	14	this is	13	and then		
2	15	i think	12	gon na	21	it 's	8	it 's	9	go to		
3	14	going to	12	in the	16	and then	8	wan na	9	name is		
4	13	it 's	11	it 's	15	do n't	8	when they	7	my name		
5	13	that 's	10	i do	14	going to	7	have to	7	so he		
6	11	to the	9	that 's	14	right here	7	here 's	6	do n't		
7	10	and he	9	there 's	14	to go	6	i know	6	he said		
8	10	ca n't	8	did n't	13	and she	6	there 's	6	he was		
9	10	have to	8	this is	13	have to	6	you know	6	i 'd		
10	9	and the	8	uh huh	13	i 'm	5	's the	6	me of		

2.5. 英語品詞使用と発達の推移

品詞使用の分布を調査することは、学習者の目標言語における文法発達を調べるのに最も基本的な分析方法の一つである。例として Kimura (2004)は、日本人英語学習者の話し言葉に見られる品詞分布を調査し、熟達度が上がるにつれ「名詞→動詞・形容詞→副詞」の順に使用が増えていくことを発見し、学習者の文構造が発達していく過程の一側面を明らかにした。

小池コーパスは CLAWS C7 のタグを用いて品詞タグ付けが行われており、一部は目視での品詞タグ修正がなされたデータである。

1~9ヶ月時点までの発話は、J、N、S3名ともインタビューアーの問いかけに対する YES, NO の受け答え、及び日本語の使用が中心であり、名詞の割合が極端に高いのは、これらの日本語使用が名詞扱いにされているためである。本来、CLAWS のタグセットには英語以外の言語に対する FW (foreign word) というタグが用意されているが、データセットでは名詞に分類されているため、本報告で特にタグを変更することはしなかった。

開始時点での名詞と形容詞の比率が高く、その後緩やかに割合が下がっていくのは、名詞や形容詞が使用されなくなるというのではなく、他の品詞の使用が増えることによる相対的なものであると思われる。確かに、6ヶ月程度までの発話の大部分を占める、動詞を含まない単語の羅列を特徴とする発話では、内容語の中でも特に意味要素の濃い名詞や形容詞の割合が高くなるのは自然である。他の品詞を見ても、限定詞、前置詞、接続詞といった機能語の割合が増えていることが見て取れる。また、動詞の使用に関しても、20%以下から始まるものの、発話が文中心になってからは約3割が動詞で安定した量が使用されている。18ヶ月以降は、3名とも動詞は依然微増傾向にあるものの、他の品詞は大きな推移は見せていない。

こういった限定詞や前置詞といった機能語の増加と、名詞や動詞といった内容語数の頭打ち感、単語レベルに留まらない英文構造の複雑化を示唆していることが推測可能である。

たとえば Kimura (2004)では、名詞が突出して多い時期から動詞が増加し、その後名詞を修飾する形容詞が増加、そして副詞が増加していく傾向を見せているものの、一貫して形容詞の数が副詞を上回っていた。一貫して形容詞の数が多という結果は、母語話者と学習者の英語を比較した Granger & Rayson (1998)においても報告されている。一方、書き言葉である JEFLL コーパスを用いた研究では、投野 (2007) では、学年が進むにつれ副詞が形容詞を上回る、小池コーパスと同様の現象が報告されている。解明には今後の研究が待たれる。

表 17 話者 N の品詞使用の推移

月数	m1-3	m4-6	m7-9	m10-12	m13-15	m16-18	m19-21	m22-24	m25-28	m29-31	m32-34
名詞	65.3%	35.6%	33.0%	27.9%	27.5%	24.0%	21.8%	20.0%	14.9%	11.7%	16.7%
動詞	14.8%	26.8%	22.4%	23.5%	20.0%	22.0%	24.2%	29.2%	31.2%	26.7%	23.4%
形容詞	8.7%	2.4%	5.2%	5.3%	4.7%	3.7%	3.3%	3.2%	3.0%	3.3%	3.2%
副詞	4.6%	3.4%	3.8%	7.4%	3.9%	6.3%	7.4%	10.6%	4.5%	11.7%	10.5%
代名詞	3.3%	19.5%	15.5%	12.3%	10.0%	10.4%	12.0%	12.9%	23.0%	10.0%	14.9%
限定詞	2.0%	4.3%	9.5%	10.3%	13.5%	16.6%	14.6%	12.4%	12.3%	20.0%	10.7%
助動詞	0.3%	1.1%	1.3%	1.6%	0.5%	0.1%	1.7%	1.1%	2.2%	3.3%	1.7%
前置詞	0.3%	4.3%	4.9%	3.9%	6.4%	5.9%	6.3%	5.2%	4.8%	5.0%	3.2%
接続詞	0.8%	2.6%	4.4%	7.9%	13.4%	11.2%	8.8%	5.4%	4.1%	8.3%	15.7%

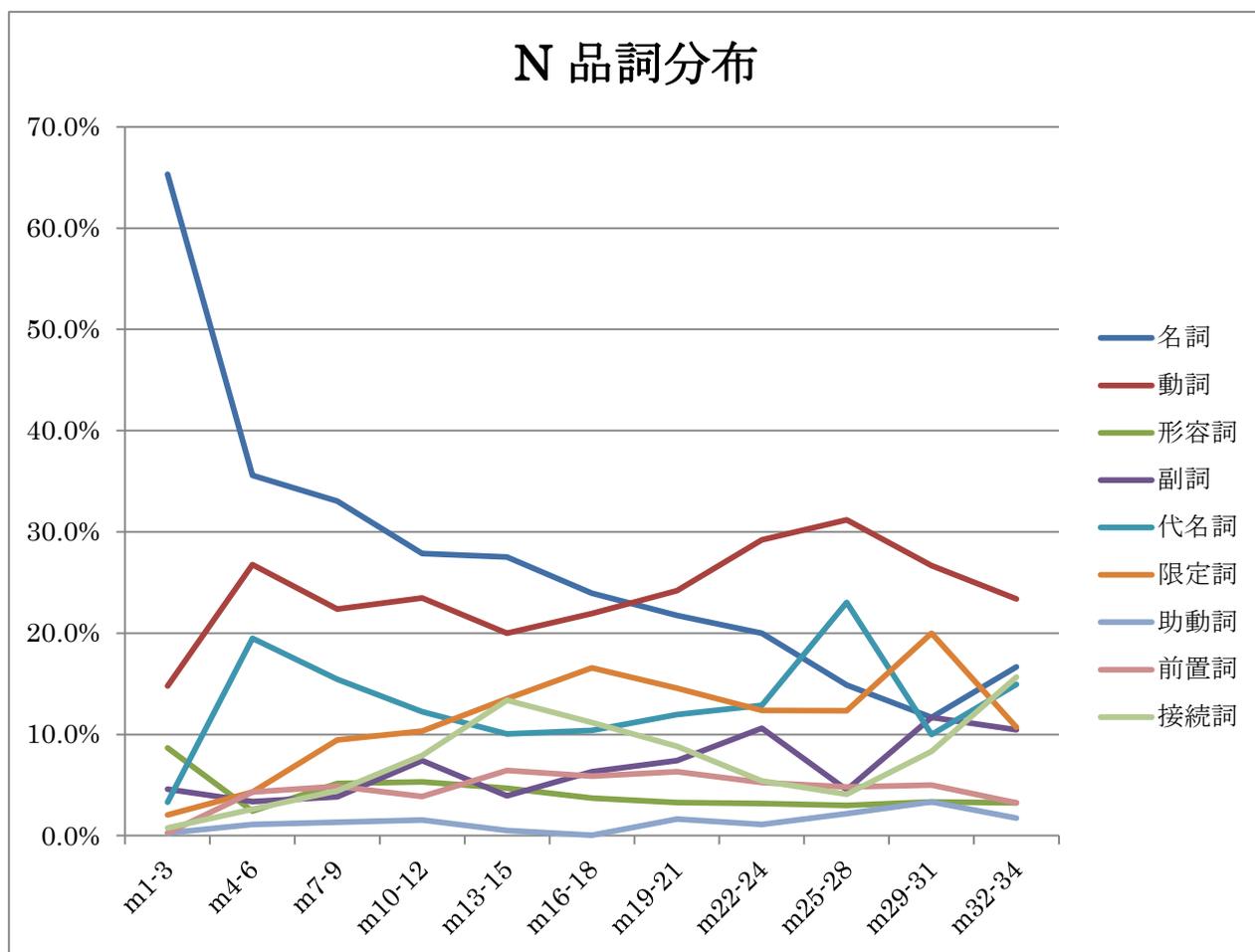


図 2 話者 N の品詞使用の推移

表 18 話者 J の品詞使用の推移

	m1-3	m4-6	m7-9	m10-12	m13-15	m16-18	m19-21	m22-24	m25-28	m29-31	m32-34
名詞	56.6%	37.6%	36.2%	29.2%	23.9%	21.5%	21.5%	20.7%	16.8%	11.9%	19.8%
動詞	11.6%	23.7%	23.6%	25.4%	23.7%	27.9%	28.0%	30.5%	29.0%	32.2%	28.7%
形容詞	24.9%	9.8%	3.9%	6.7%	4.5%	3.7%	3.8%	2.3%	3.1%	3.9%	4.1%
副詞	1.4%	1.9%	2.7%	7.6%	7.4%	10.2%	7.5%	9.2%	10.8%	12.5%	10.4%
代名詞	1.9%	12.9%	13.6%	6.2%	14.1%	15.2%	16.3%	19.2%	19.0%	16.6%	18.7%
限定詞	1.9%	6.4%	6.1%	5.9%	6.9%	5.4%	5.8%	5.1%	5.3%	9.2%	5.1%
助動詞	0.2%	1.5%	1.2%	2.0%	1.5%	2.1%	2.1%	1.7%	3.3%	2.6%	2.9%
前置詞	0.5%	2.3%	3.6%	4.7%	6.5%	5.8%	6.1%	6.0%	5.4%	4.9%	5.1%
接続詞	0.9%	4.1%	9.2%	12.4%	11.5%	8.2%	8.8%	5.4%	7.3%	6.3%	5.2%

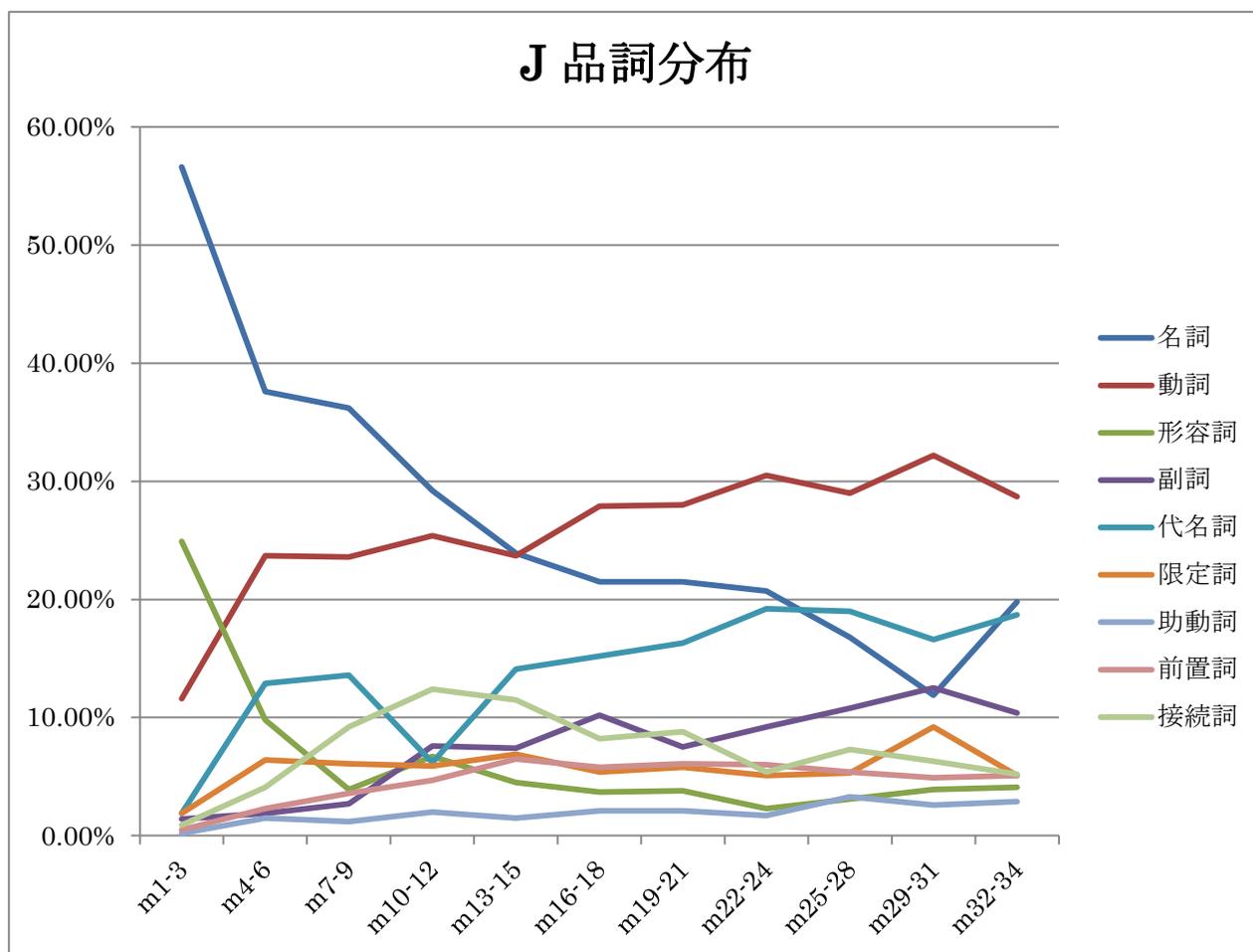


図 3 話者 J の品詞使用の推移

表 19 話者 S の品詞使用の推移

	m1-3	m4-6	m7-9	m10-12	m13-15	m16-18	m19-21	m22-24	m25-28	m29-31	m32-34
名詞	53.5%	23.5%	18.8%	14.7%	22.4%	18.1%	19.8%	18.7%	13.3%	10.7%	18.3%
動詞	16.3%	24.0%	25.3%	26.3%	22.9%	26.4%	26.2%	27.9%	27.5%	27.6%	24.9%
形容詞	9.4%	2.9%	3.1%	3.0%	3.4%	2.9%	3.6%	3.2%	3.4%	3.6%	4.1%
副詞	5.2%	14.1%	10.7%	10.3%	8.5%	9.5%	9.3%	10.5%	13.4%	12.7%	10.0%
代名詞	3.1%	17.0%	19.4%	20.0%	13.5%	17.0%	15.9%	16.1%	18.2%	20.5%	18.8%
限定詞	3.1%	11.2%	9.9%	11.8%	15.2%	12.6%	12.4%	11.9%	10.4%	11.5%	9.5%
助動詞	0.0%	1.3%	2.8%	2.6%	1.3%	2.4%	2.1%	2.0%	3.5%	4.0%	3.0%
前置詞	9.4%	2.9%	3.1%	3.0%	3.4%	2.9%	3.6%	3.2%	3.4%	3.6%	4.1%
接続詞	0.0%	3.1%	6.9%	8.5%	9.5%	8.1%	7.4%	6.7%	6.9%	5.9%	7.3%

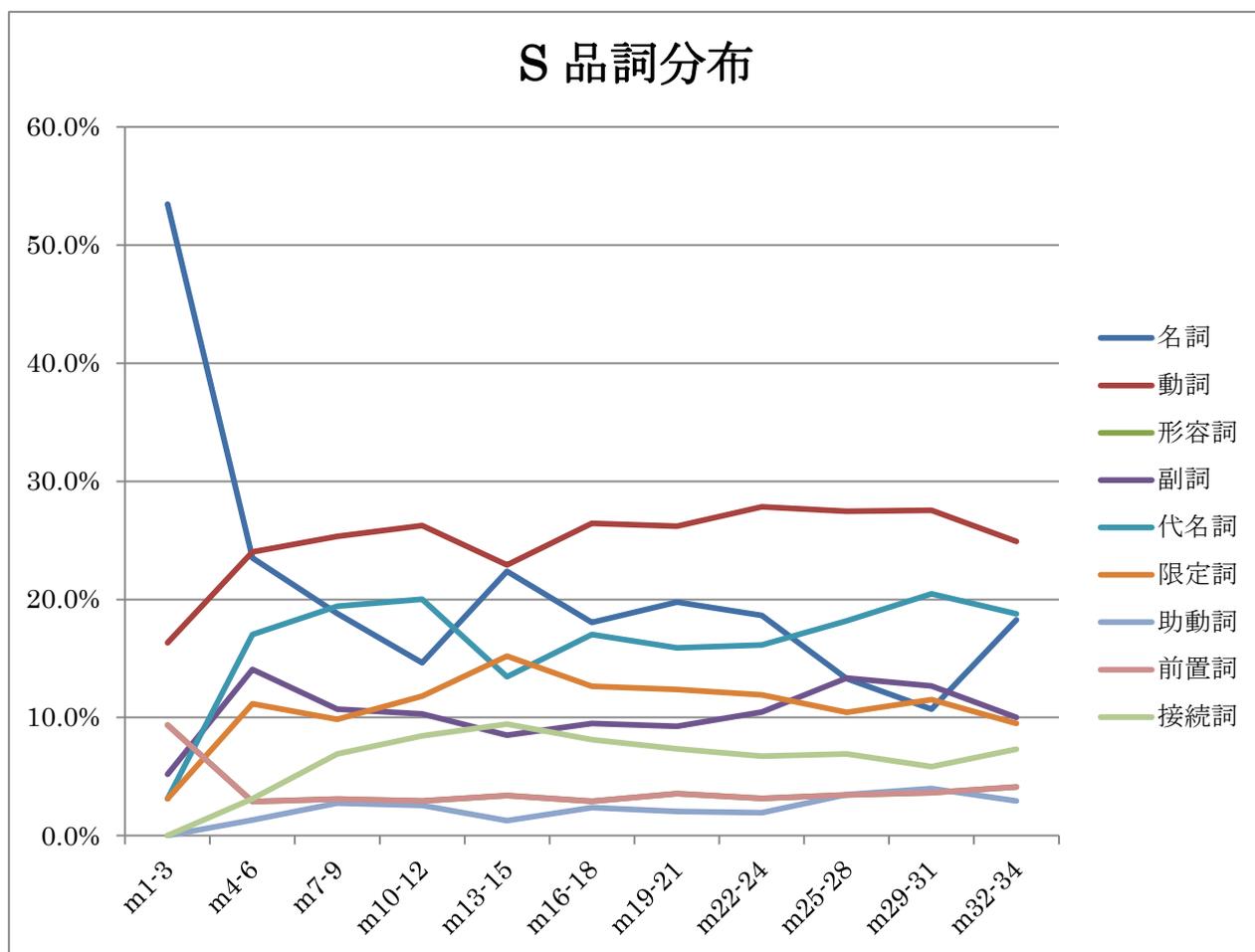


図 4 話者 J の品詞使用の推移

また、本項では品詞データを用いてのクラスター分析をおこなった。

クラスター分析では、距離行列を利用して似た傾向を持つグループをまとめることができる。図 5

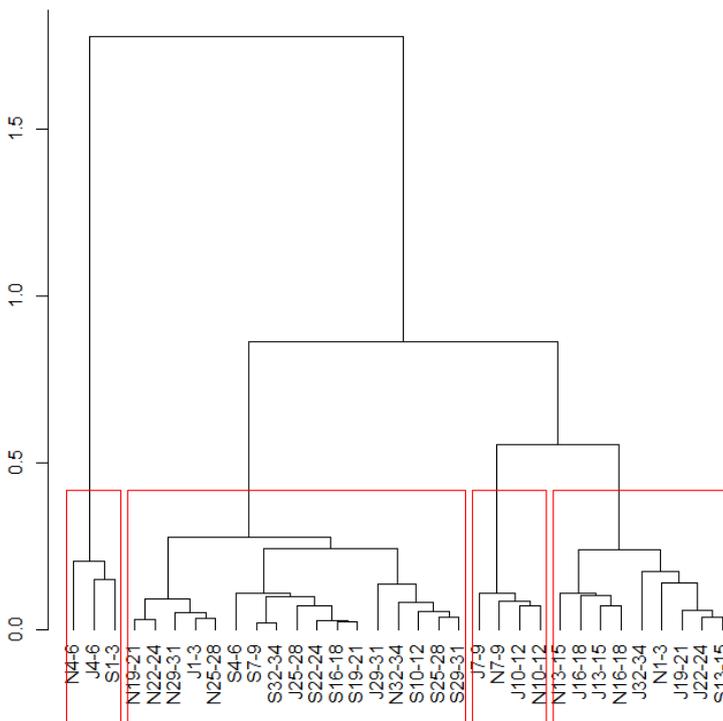
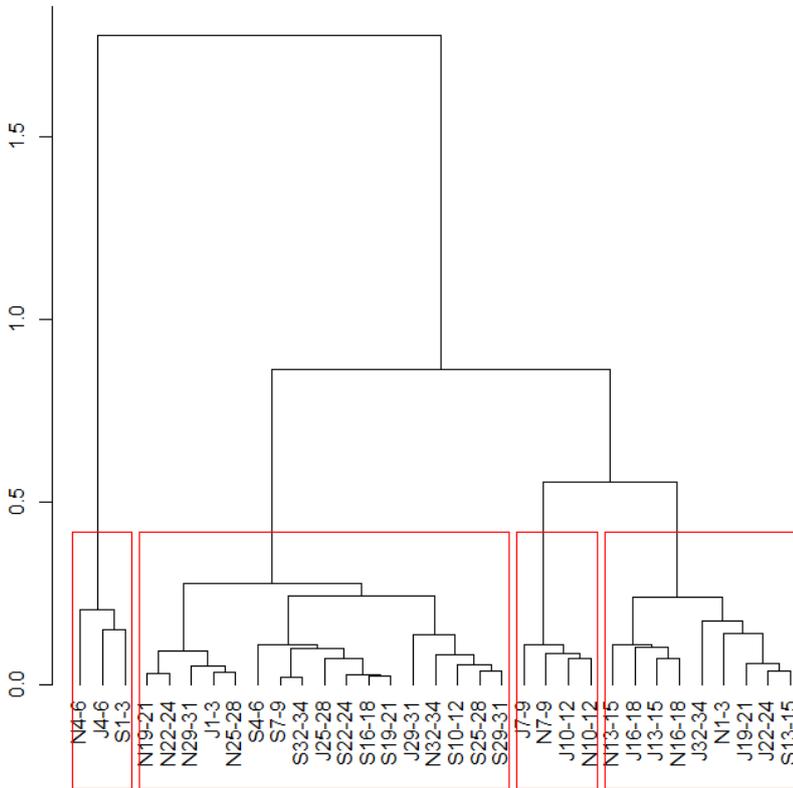


図 6 は、正規化した品詞の頻度を用いたクラスター分析の結果である。この散布図及びデンドログラムによると、まず大別して、N, J の 4~6 ヶ月目、及び S の 1~3 ヶ月目と、それ以外の発話データで大分できるようである。これは最初期のグループとみなすことができるだろう。また、13 ヶ月~18 ヶ月のデータも似た傾向にあるようである。

散布図、デンドログラムを観察すると、S のデータが N や J とは異質である可能性が示唆される。デンドログラムで最も大きいグループは、7 ヶ月以降の S の発話に、N や J の 24 ヶ月以降のものが混じる形となっている。N と J の発話は似た傾向を示しており、事実散布図やデンドログラムでも近いところにあるが、S の発話とは特殊であるといえるようだ。

N、J と比べて爆発的に発話量が増加した時点が早かった、S の語彙発達の結果と合わせてみると、渡航時最年少だった S は、最も英語の習得が早かったといえるかもしれない。特にクラスター分析において、他 2 人の渡航時最終年と似た傾向を示しているのは、S がより早くその熟達段階に達していた可能性を示唆している。

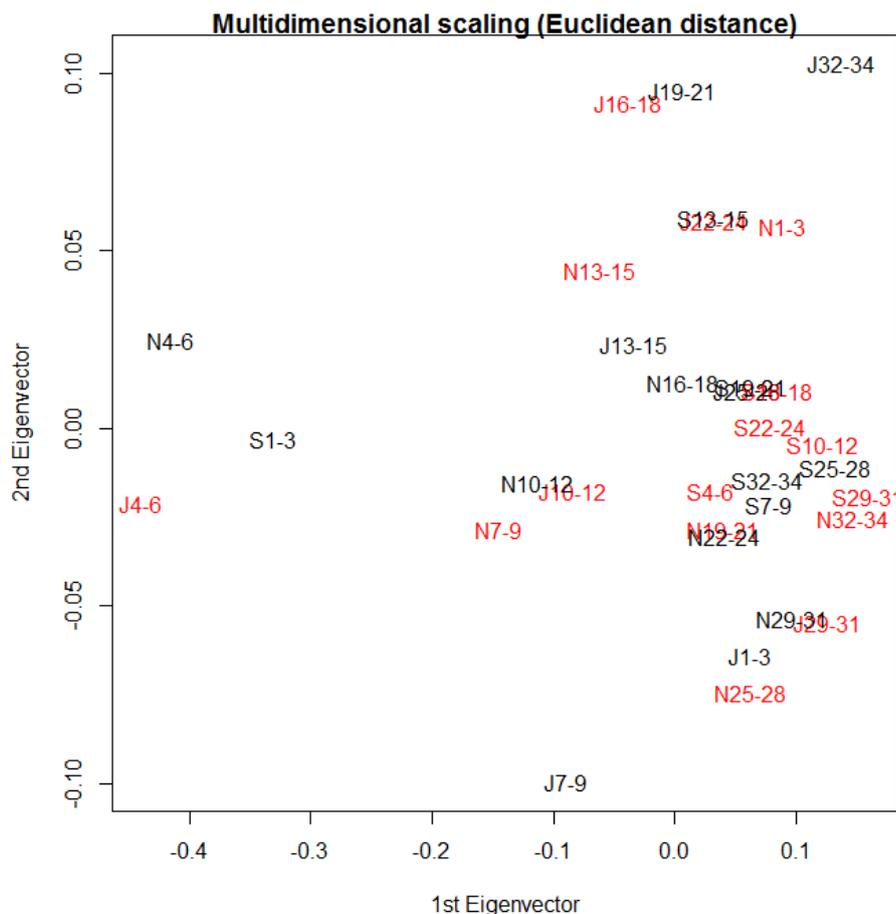


図 5 品詞頻度によるクラスター分析 (ユークリッド距離を用いた MDS 散布図)

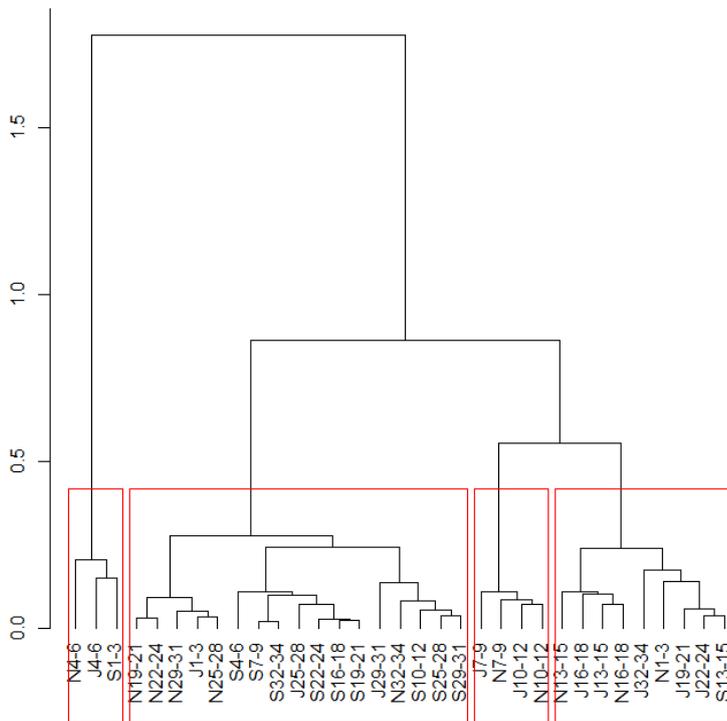


図 6 品詞による階層的クラスター分析

2.6. 品詞連鎖の推移

まず最初期から現れる **nn1** の連鎖は、一般名詞の単数形である。これは前述の通り、日本語もしくはカタカナ語として使用されている英単語に **nn1** のタグが振られていることに起因する。一部のデータでは日本語の単語に **jp** という **CLAWS C7** にない独自のタグが振られており、これも 1~3 ヶ月の間などは上位にあることが確認できる。**uh uh** の連鎖も上位に入っているが、これは間投詞に **yes/no** が含まれているため、特に最初期はインタビューアーの質問に対して **yes, no** の返答のみが多いために起きている連鎖である。

発話量が増加してから、ほぼ常に 1 位である **at nn1** というのは、冠詞+名詞の組み合わせである。**at** は中でも、**the** や **no** に付与される定冠詞で、**a, an, every** などの不定冠詞 **at1** と区別される。日本人英語学習者を研究した他の研究の中で指摘され、小池コーパスにみられない点として、冠詞の使用が多いこと、一人称代名詞の使用が少ないことがあげられる。Tono (1999) は、**JEFLL** コーパスの中 1~中 3、高 2 のデータに大学生の作文データを加えて **3-gram** を抽出し、**Aarts & Granger (1998)** の研究との比較を行った。結果、習熟度に従って名詞中心の連鎖から動詞中心の連鎖へとシフトすること。助動詞、冠詞の出現が極めて少ないことを報告している。小池コーパスにおいても、動詞の増加は確認されたが、前述のように冠詞の出現率は決して低くない。3 名いずれも、12 ヶ月以降は定冠詞+名詞の組み合わせが必ず上位に入っている。これは、米国滞在時ゆえの学習環境、学習内容の違いによる影響であると考えられるが、母語として日本語を共有する学習者間の比較は非常に興味深い。日本人学習者による冠詞の過小使用はしばしば指摘されるところであるが、熟達度が上がるにつれ冠詞の使用頻度は多くなっていくこともまた知られており、学習環境による影響で急速に英語力を身につけられているとすれば、教育的には大きな示唆となる可能性を秘めているだろう。

表 20 N の品詞 bigram 上位 10 位の推移

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	141	nn1 nn1	78	jp jp	183	jp jp	95	at nn1	177	at nn1	102	at nn1
2	44	uh uh	36	uh uh	32	ppis1 vv0	89	jj nn1	165	nn1 cc	73	nn1 cc
3	41	mc mc	27	ppis1 vv0	31	at nn1	77	nn1 nn1	64	ii at	37	cc at
4	25	jp jp	21	nn1 ppis1	30	jj nn1	64	nn1 cc	64	nn2 cc	35	nn1 vvd
5	13	nn1 vv0	15	xx vvi	28	appge nn1	64	nn1 vbz	60	cc pphs1	31	ii at
6	10	np1 np1	14	np1 np1	24	uh uh	42	ii at	58	nn1 vbz	30	to vvi
7	10	vv0 vv0	14	vv0 nn1	21	nn1 ppis1	42	nn1 vvd	57	jj nn1	26	jj nn1
8	9	jj jj	13	nn1 nn1	19	dd1 vbz	38	xx vvi	47	cc at	26	pphs1 vvd
9	9	uh jj	13	nn1 uh	19	nn1 jp	37	appge nn1	46	pphs1 vvd	25	appge nn1
10	9	uh nn1	13	ppis1 vd0	19	nn1 uh	36	cc nn1	45	cc nn1	24	dd1 vbz
月数	21		24		28		31		34			
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖		
1	79	at nn1	23	at nn1	37	at nn1	2	dd1 pn1	19	pphs1 vvd		
2	48	ii at	14	xx vvi	22	ii at	2	pn1 vbz	17	nn1 cc		
3	42	nn1 cc	13	vd0 xx	21	to vvi	2	vdz xx	15	cc rt		
4	42	to vvi	12	jj nn1	21	xx vvi	2	vm vvi	14	at nn1		
5	35	at1 nn1	12	to vvi	18	ppis1 vv0	2	vvi cc	10	to vvi		
6	33	nn1 nn1	10	ii at	16	nn1 ii	2	vvi rp	9	nn1 vvd		
7	26	appge nn1	10	nn1 nn1	16	uh ppis1	2	xx vvi	8	cc at		
8	26	at jj	10	vbz at1	16	vm vvi	1	appge nn1	8	cc pphs1		
9	26	jj nn1	9	ex vbz	15	ppis1 vvd	1	at nn1	7	vvd rp		
10	23	jp jp	9	nn1 uh	13	nn1 uh	1	cc at	6	np1 vvd		

表 21 J の品詞 bigram 上位 10 位の推移

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	143	nn1 nn1	28	nn1 nn1	91	jp jp	128	at nn1	264	at nn1	121	at nn1
2	51	jp jp	25	mc mc	37	nn1 cc	122	nn1 cc	261	nn1 cc	70	nn1 cc
3	33	uh uh	20	jp jp	32	nn1 uh	97	jj nn1	109	at1 nn1	58	at1 nn1
4	17	mc mc	19	vv0 nn1	30	dd1 vbz	59	nn1 vvd	105	jj nn1	46	xx vvi
5	16	jj jj	18	uh uh	29	nn1 nn1	43	cc at	103	ii at	44	ex vbz
6	12	nn1 jj	12	uh nn1	27	appge nn1	41	appge nn1	88	nn1 nn1	42	nn1 nn1
7	12	nn1 vv0	11	nn1 cc	26	at nn1	41	ii at	87	appge nn1	41	ii at
8	10	vv0 vv0	9	nn1 uh	26	nn1 vbz	39	xx vvi	81	dd1 vbz	41	to vvi
9	9	jj nn1	8	ppy vvi	24	uh uh	35	nn1 uh	77	cc pphs1	37	appge nn1
10	9	np1 np1	7	mc uh	19	jj nn1	34	dd1 vbz	73	cc at	36	dd1 vbz
月数	21		24		28		31		34			
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖		
1	131	at nn1	41	at nn1	160	at nn1	23	at nn1	26	to vvi		
2	55	nn1 cc	19	xx vvi	71	to vvi	18	to vvi	22	appge nn1		
3	50	ii at	18	nn1 nn1	58	ii at	11	vm vvi	22	jj nn1		
4	42	at1 nn1	16	vd0 xx	54	nn1 cc	10	dd1 vbz	18	at nn1		
5	35	to vvi	15	appge nn1	50	xx vvi	8	pph1 vbz	18	nn1 nn1		
6	32	jj nn1	15	to vvi	41	nn1 nn1	7	at nn2	18	pphs1 vvd		
7	31	cc at	12	ii at	40	vm vvi	7	dd1 nn1	18	ppy vvi		
8	31	ppis1 vv0	11	ppis1 vd0	38	at1 nn1	7	ppis1 vv0	15	vm ppy		
9	23	nn1 nn1	10	uh uh	38	ppis1 vv0	7	vbz at	13	at1 nn1		
10	19	at jj	9	nn1 ii	37	dd1 vbz	7	vvgk to	13	nn1 cc		

表 22 S の品詞 bigram 上位 10 位の推移

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	80	nn1 nn1	54	uh uh	107	jp jp	109	nn1 cc	305	at nn1	96	at nn1
2	28	uh uh	30	nn1 nn1	73	dd1 vbz	107	at nn1	190	nn1 cc	60	to vvi
3	17	mc mc	23	dd1 vbz	64	vd0 xx	85	to vvi	129	at1 nn1	55	nn1 cc
4	11	nn1 uh	19	nn1 uh	64	xx vvi	66	at1 nn1	126	ii at	52	appge nn1
5	11	vv0 vv0	19	nn1 vv0	55	appge nn1	62	ii dd1	87	ex vbz	51	jj nn1
6	10	uh nn1	18	ppis1 vv0	55	uh uh	59	xx vvi	82	vbz vvg	49	at1 nn1
7	9	jj jj	18	uh nn1	54	ppis1 vv0	54	appge nn1	74	cc at	45	ppis1 vv0
8	9	vv0 nn1	18	vd0 xx	54	to vvi	54	dd1 vbz	71	nn1 vbz	43	pphis1 vvd
9	8	nn1 jj	16	nn1 rr	53	ppy vm	53	ppis1 vv0	62	at nn2	41	ii at
10	8	nn1 vv0	15	vv0 nn1	43	vm xx	51	ii at	62	nn1 nn1	39	nn1 vvd
月数	21		24		28		31		34			
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖		
1	87	at nn1	43	at nn1	80	at nn1	25	at nn1	38	at nn1		
2	42	nn1 cc	38	at1 nn1	68	to vvi	24	to vvi	30	pphis1 vvd		
3	40	at1 nn1	34	to vvi	39	ppis1 vm	18	ppis1 vv0	26	jj nn1		
4	40	ii at	31	uh uh	36	ii at	18	vm vvi	26	to vvi		
5	33	to vvi	25	xx vvi	30	uh uh	17	dd1 vbz	17	nn1 cc		
6	32	ppis1 vv0	23	ii at	30	vm vvi	17	pphis2 vv0	15	cs pphis1		
7	31	jj nn1	21	ppis1 vv0	29	ppis1 vv0	15	ppy vv0	13	cc rt		
8	30	np1 vvz	20	nn1 uh	26	xx vvi	13	appge nn1	13	ii at		
9	27	pphis1 vvd	18	nn1 cc	22	nn1 cc	13	xx vvi	13	vm vvi		
10	27	xx vvi	18	vd0 xx	22	pph1 vbz	10	at1 nn1	12	appge nn1		

2.7. データの抽出について

本報告で用いたデータは、小池コーパスには、本文のほかにヘッダー、話者タグ、センテンスタグ、コメントなどが付与されている。本来であれば、タグを利用することで自動的に学習者ごとのデータを抽出が可能だが、一部タグに不具合・不整合が確認されたため、分析に至るまでにデータの下処理を行っている。

不具合・不整合が存在し、また規則的な不具合でもなかったため、本報告で用いたデータは消去法を用いている。即ち、対象となっている学習者以外の発話やメタ情報を消していくことで、最終的に目的の学習者の発話のみを抽出することが目的である。

第一段階として、まず話者タグ、センテンスタグに基づいて発話を除去した。分析対象がJだとすると、J以外の話者の発言を自動消去する作業となる。しかし、タグの不整合がある個所については除去されず残るため、2名で目視によるデータのクリーニングを行った。目視による除去が終わった後、コーパスのタグを全て抽出し、対象の学習者以外のタグが混じっていないかダブルチェックを行い、全ファイルの内容が対象のみになったことを確認した上で、センテンス・話者タグを全て消去した。

本報告で使用したデータは、以上のプロセスによって得られたデータである。

次に、いくつかデータの計算方法などについて、補遺を記す。

【語彙について】

- ・ ハイフン、アンダーバー、アポストロフィーは英語の通例に従って単語の一部とみなしたが、データにはその他コロン・セミコロン（日本語の長音表記としての使用）など、予期しない使用が散見されるため、厳密には修復の必要がある。ただし、長音表記が必ずしもコロン・セミコロンなどによって表現されているわけではない。おそらく記述者間のブレが原因である。
- ・ afraid に対して、'fraid など音表記が行われている個所があり、おそらく統一はされていない。
- ・ 発話の日本語部分には、独自の JP タグが付与されている個所と、CLAWS の出力と思われる名詞が付与されている個所がある。このため、JEFLC コーパスなどで行われている、日本語部分のみを抽出した分析は現状不可能である。
- ・ 品詞リストにノイズが含まれるものがあるが、これは仕様である。通常、テキスト内の数字は MC 系のタグで表現されるはずだが、n_56 のように、数字がタグ位置に挿入されているものが混じっている。特定後に処理することが可能。

【タグ修正について】

- ・ POS タグ修正という記録が残っているため、推測するに「タグ修正」という報告は品詞タグの修正を行ったのは意味だと思われる。実際、CLAWS7 tagset には存在しない JP（日本語の単語）というタグがついている個所がある。しかし JP タグも、英語からの借用語として日本語に存在するカタカナ語に関してはついていないなど、網羅的なわけではない。

【話者タグについて】

<J> 2842 件
</J> 2831 件
<N> 1779 件
</N> 1773 件
<S> 3444 件
</S> 3439 件
<s> 23705 件
</s> 23706 件

- ・ズレは数件に見えるが、閉じタグ（開始）→閉じタグ（終了）などが混在しているため、実際のズレはより数が多い。Typo と思われる箇所も。
- ・その他、ヘッダーに関しても正確にはタグ付けがなされていない。たとえばあるファイルでは transcriber が書き起こした人なのに対して、comp_transcriber タグが使用されているファイルがあるなど。
- ・他、<head></head>、<body></body>の数も一致しない。

参考文献

Aarts, J., & Granger, S. (1998). Tag sequences in learner corpora: A key to interlanguage grammar and discourse. In S. Granger (Ed.), *Learner English on Computer* (pp. 132-141). Harlow: Longman.

Kimura, M. (2004). How do Japanese EFL learners develop their interlanguage grammar?: Applying the N-gram model in seeking L2 developmental indices. *Handbook of an International Symposium on Learner Corpora in Asia* (pp. 53-58).

Tono, Y. (1999). A corpus-based analysis of interlanguage development: Analysing part-of-speech tag sequences of EFL learner corpora. In B. Lewandowska-Tomaszczyk, & P.J. Melia (Eds.), *PALC '99 Practical Applications in Language Corpora* (pp. 323-339). Frankfurt: Peter Lang.

シャノン C E・ヴィーヴァー W. (長谷川淳・井上光洋訳) (1964). 『コミュニケーションの数学的理論』東京：明治図書

投野由紀夫. (編著) (2007) 『日本人中高生 1 万人の英語コーパス』東京：小学館

付録 A. 話者 N 基礎語彙表

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	39	no	58	i	65	i	174	and	332	and	167	the
2	20	yes	39	no	54	no	136	the	209	the	142	and
3	11	kam	39	yes	51	and	111	is	116	is	53	to
4	11	tip	22	do	49	the	97	i	85	he	45	this
5	9	do	21	you	37	to	80	's	72	's	43	's
6	7	gud	15	n't	29	you	67	this	59	this	37	is
7	7	hi	15	to	27	's	56	no	53	on	34	wolf
8	7	ho	12	know	26	my	55	to	51	there	31	house
9	6	k	12	like	25	is	51	n't	49	i	26	so
10	6	to	11	and	22	do	47	he	38	a	25	he
11	5	a	11	english	21	a	44	do	38	boy	25	that
12	5	bai	11	go	20	yes	44	yes	35	has	22	it
13	5	hia	9	a	18	this	44	you	30	one	22	went
14	5	ta	9	play	17	go	38	that	30	to	20	i
15	5	ten	9	school	16	in	36	said	26	here	20	pig
16	4	ba	8	can	16	like	35	it	26	pinocchio	20	you
17	4	haus	8	in	16	tte	35	then	25	said	19	was
18	4	n	8	my	15	he	33	here	24	that	18	mother
19	4	tu	7	the	15	it	32	a	23	in	18	we
20	3	ai	7	tte	15	not	32	so	23	put	17	a

続き

月数	21		24		28		31		34	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	162	the	35	the	68	i	4	this	49	and
2	105	and	27	i	55	the	3	that	31	the
3	73	to	25	and	38	you	2	and	24	he
4	50	that	23	to	37	to	2	but	18	then
5	49	i	22	's	35	's	2	does	15	to
6	45	a	22	n't	35	it	2	go	12	said
7	43	's	18	a	30	n't	2	is	11	pinocchio
8	37	gulliver	18	do	27	that	2	n't	9	i
9	34	so	18	is	23	we	2	one	7	a
10	30	n't	15	you	21	have	2	some	7	legs
11	29	in	14	here	20	do	2	what	6	father
12	29	is	12	it	17	and	1	'm	6	on
13	28	have	12	not	17	know	1	's	6	so
14	28	you	12	this	16	no	1	about	6	was
15	27	no	11	there	16	they	1	back	5	him
16	27	there	10	car	16	this	1	better	5	his
17	25	he	10	school	15	a	1	can	5	put
18	25	we	9	have	15	oh	1	cause	5	went
19	24	go	8	no	12	'm	1	come	4	's
20	21	was	7	are	12	go	1	hard	4	head

付録 B. 話者 J 基礎語彙表

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	39	no	33	yes	68	and	244	and	473	and	213	the
2	15	yes	23	no	56	is	174	the	353	the	155	and
3	7	:	21	is	51	yes	110	you	192	's	138	's
4	7	to	20	and	49	no	75	i	176	this	88	there
5	6	do	19	i	38	i	74	's	161	is	87	a
6	6	kam	18	you	34	the	72	is	155	a	80	you
7	4	and	13	english	32	my	51	n't	145	he	75	i
8	4	buk	13	have	32	this	48	little	104	to	74	is
9	4	n't	12	do	31	's	45	do	102	you	68	it
10	4	onnanoko	9	one	27	n't	45	yes	92	it	68	n't
11	4	pardon	8	know	23	we	44	that	90	i	64	this
12	4	white	8	seven	21	do	44	to	85	in	53	do
13	4	yah	8	speak	19	that	42	black	77	there	53	to
14	3	ai	7	n't	19	to	41	no	71	that	48	no
15	3	bat	7	two	14	name	40	it	66	was	44	here
16	3	brown	6	can	13	know	37	said	65	like	42	he
17	3	gu	6	five	13	you	36	this	58	here	39	have
18	3	i	6	me	12	like	35	sambo	56	her	36	of
19	3	iuno	6	tape	11	a	34	tiger	56	n't	36	one
20	3	kandi	6	three	11	one	32	one	52	but	36	that

続き

月数	21		24		28		31		34	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	178	the	54	the	180	the	40	the	57	you
2	121	and	36	n't	118	and	28	this	45	to
3	79	i	33	i	114	i	24	it	32	and
4	72	to	32	you	108	's	23	's	32	the
5	64	's	28	it	90	it	23	they	30	he
6	55	a	26	's	81	you	20	and	29	a
7	55	this	25	do	79	they	19	here	27	's
8	50	you	25	to	67	she	19	is	19	okay
9	47	it	17	and	67	to	18	have	18	it
10	41	he	16	is	58	n't	18	to	17	like
11	37	is	14	he	55	have	16	i	16	i
12	36	in	13	have	52	this	15	that	15	man
13	36	n't	12	this	48	a	14	we	15	so
14	34	have	11	not	41	is	12	in	15	would
15	31	was	11	yes	40	in	12	you	14	do
16	26	there	10	a	40	pretend	11	one	14	oh
17	26	they	10	on	39	do	9	if	14	some
18	25	are	10	that	38	that	9	n't	13	hi
19	25	that	10	your	36	we	9	okay	13	n't
20	25	what	9	are	33	here	9	put	12	go

付録 C 話者 S 基礎語彙表

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	32	no	48	this	190	you	307	and	415	and	167	the
2	18	yes	46	you	181	i	194	i	412	the	164	and
3	5	dog	41	i	169	this	180	's	196	's	109	's
4	5	du	39	no	160	and	178	it	168	a	107	i
5	5	pardon	33	is	134	do	177	we	142	is	97	you
6	4	do	32	and	130	it	175	the	118	to	86	it
7	4	faib	29	k	115	n't	171	you	109	there	83	to
8	4	gu	29	o	98	to	163	that	104	i	71	a
9	4	kam	27	yes	93	is	151	this	92	it	69	that
10	4	n	24	do	89	's	127	to	92	you	64	she
11	3	a	23	that	86	no	103	do	88	this	56	then
12	3	ap	21	like	73	that	102	n't	84	on	53	n't
13	3	dadi	21	n't	59	like	91	a	83	she	53	said
14	3	dentist	21	the	56	the	85	have	75	then	51	this
15	3	in	19	's	55	here	85	like	69	that	49	is
16	3	mo	19	monkey	53	o	70	is	59	in	48	no
17	3	neim	16	everybody	52	have	69	no	55	he	40	in
18	3	orend	15	come	52	we	65	he	54	we	38	do
19	3	sit	15	now	49	k	61	're	51	her	34	little
20	3	vest	15	to	39	ca	58	she	44	dick	33	there

続き

月数	21		24		28		31		34	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	128	the	61	and	145	the	43	it	56	the
2	111	and	57	the	108	and	40	the	51	and
3	105	i	54	you	106	i	40	they	46	i
4	63	's	52	a	92	to	39	and	45	he
5	62	a	51	's	71	you	39	i	45	to
6	61	to	50	i	66	's	35	you	30	you
7	55	that	43	to	66	she	31	's	19	go
8	47	it	41	is	63	it	31	this	18	like
9	47	she	39	n't	48	is	24	is	18	was
10	42	in	37	it	45	that	21	that	17	a
11	42	n't	35	now	45	we	20	here	17	we
12	41	is	23	do	42	here	20	we	16	is
13	41	you	22	that	41	n't	18	on	16	it
14	37	he	21	in	38	okay	17	n't	16	me
15	32	means	21	well	37	they	16	a	16	said
16	28	one	20	get	36	this	16	see	15	on
17	26	have	19	there	34	a	15	to	15	so
18	26	there	18	this	34	one	14	like	15	then
19	25	this	17	he	31	do	14	too	14	have
20	23	was	17	what	30	'll	13	have	13	's

付録 D. N の bigram 使用上位 10 位

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	16	no no	13	i do	13	do you	33	do n't	51	and he	37	and the
2	10	yes no	13	yes yes	10	you like	20	this is	44	and the	28	the wolf
3	9	no yes	12	do n't	9	that 's	19	little black	32	on the	20	this is
4	5	hi tip	10	n't know	9	this is	17	black sambo	25	there is	18	house and
5	5	tip hi	9	do you	8	go to	17	pulling the	19	and there	14	is the
6	4	kam hia	9	yes no	8	no i	16	gon na	14	and white	14	went to
7	3	gud bai	7	go to	6	do n't	15	and he	14	in the	12	and he
8	3	nain ten	7	no yes	6	i am	14	i do	14	there 's	9	's house
9	3	wan tu	6	i go	6	i do	14	is pulling	14	this is	9	to the
10	3	za do	6	no no	6	in the	14	tiger said	13	boy is	8	wolf came
月数	21		24		28		31		34			
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖		
1	21	the gulliver	11	do n't	14	have to	2	does n't	15	and then		
2	19	and the	7	it 's	12	do n't	2	one is	8	and he		
3	16	have to	6	's a	12	i 'm	2	this one	8	and the		
4	16	in the	6	and there	11	it 's	1	'm teddy	5	then the		
5	15	the other	6	there 's	11	n't know	1	's your	4	him and		
6	12	do n't	5	they are	9	on the	1	about trudy	4	Pinocchio said		
7	12	to the	5	this is	8	in the	1	and but	4	the Pinocchio		
8	11	on the	4	go to	8	who 's	1	and the	4	then he		
9	10	other island	4	i do	7	that 's	1	back this	3	he put		
10	10	so the	4	is a	6	'm catching	1	better you	3	he went		

付録 E. J の bigram 使用上位 10 位

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	18	no no	7	do n't	19	this is	40	and the	67	and he	41	there 's
2	4	and white	7	i do	18	do n't	40	little black	57	and the	36	do n't
3	4	do n't	7	n't know	14	i do	33	black sambo	53	this is	27	and the
4	4	no yes	6	can you	12	n't know	28	do n't	43	there 's	25	it 's
5	4	yes yes	6	no yes	11	that 's	20	that 's	40	and there	20	's a
6	3	i do	6	speak english	10	yes yes	19	in the	38	in the	20	this is
7	3	n't know	6	yes yes	7	name is	18	i 'll	31	's a	18	have to
8	3	to iuno	5	do you	7	no i	17	the tiger	31	and this	16	over here
9	2	bag map	5	i have	6	i like	15	n't know	29	he 's	16	that 's
10	2	balu:n tennis	5	is coming	6	this one	14	and little	27	and then	14	over there
月数	21		24		28		31		34			
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖		
1	31	and the	16	do n't	23	have to	7	have to	14	would you		
2	17	it 's	11	i do	22	let 's	7	it 's	14	you like		
3	15	do n't	8	n't know	21	i 'll	7	this is	11	like to		
4	15	have to	7	have to	21	she 's	6	let 's	11	to say		
5	15	in the	6	's your	20	it 's	5	's the	10	and he		
6	13	to the	5	did n't	19	do n't	5	do n't	9	say something		
7	12	the barber	5	it 's	16	and then	4	going to	8	i 'm		
8	11	and then	5	n't have	16	gon na	4	gon na	8	it 's		
9	11	i know	5	the fire	16	just pretend	4	n't have	8	thank you		
10	10	on the	5	you have	16	this is	4	the kitchen	8	to go		

付録 F. S の bigram 使用上位 10 位

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	15	no no	29	o k	65	do n't	50	do n't	82	there 's	26	that 's
2	6	no yes	18	do n't	57	this is	49	it 's	73	and the	24	it 's
3	4	yes no	15	this is	49	o k	41	gon na	59	and then	23	do n't
4	3	no pardon	12	no no	40	have to	41	we 're	57	on the	23	then the
5	2	ap do	11	ma ma	39	ca n't	38	like this	45	's a	22	i think
6	2	berj gud	8	for you	27	like this	35	have to	45	and there	19	she said
7	2	boj grl	8	yes yes	26	it 's	34	and then	26	in the	18	in the
8	2	dadi gu	7	round and	25	i do	29	that 's	25	's the	15	and the
9	2	du ju	7	that 's	24	do it	26	ca n't	22	this is	15	the little
10	2	du zat	6	is for	21	you ca	25	and the	20	and he	14	then she
月数	21		24		28		31		34			
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖		
1	17	in the	18	do n't	27	i 'll	14	this is	13	and then		
2	15	i think	12	gon na	21	it 's	8	it 's	9	go to		
3	14	going to	12	in the	16	and then	8	wan na	9	name is		
4	13	it 's	11	it 's	15	do n't	8	when they	7	my name		
5	13	that 's	10	i do	14	going to	7	have to	7	so he		
6	11	to the	9	that 's	14	right here	7	here 's	6	do n't		
7	10	and he	9	there 's	14	to go	6	i know	6	he said		
8	10	ca n't	8	did n't	13	and she	6	there 's	6	he was		
9	10	have to	8	this is	13	have to	6	you know	6	i 'd		
10	9	and the	8	uh huh	13	i 'm	5	's the	6	me of		

付録 G. N の trigram 使用上位 10 位

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	10	no no no	12	i do n't	9	do you like	33	little black sambo	7	and he put	17	and the wolf
2	5	no yes no	9	do n't know	6	i do n't	17	is pulling the	7	and there is	13	this is the
3	5	tip hi tip	6	do you like	4	go to school	14	i do n't	7	he put the	8	's house and
4	4	yes no yes	6	yes yes yes	4	tte nante iuno	12	do n't know	6	and he said	8	house and the
5	2	ai maus eapo	5	go to school	4	you like what	11	pulling the old	6	and the girl	7	the wolf came
6	2	bla hipopotms dog	3	i go to	3	a rich man	8	gon na eat	6	and there 's	6	you have to
7	2	blak sri ho	3	no i do	3	brothers and sisters	8	na eat you	6	boy and the	5	this day is
8	2	dentist polisman milkman	3	no yes yes	3	do n't know	7	eat you tiger	6	sitting on the	5	we went to
9	2	di monki taiga	3	to school tommorrow	3	my father 's	7	you tiger said	5	and he has	4	and nd pig
10	2	elehant zi bla	3	tte nante iuno	3	my room i	6	and he 's	5	and he went	4	he went to

続き

月数	21		24		28		31		34	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	10	the other island	4	and there 's	7	do n't know	2	this one is	5	and then the
2	7	i do n't	4	do n't know	6	'm catching up	1	's your name	4	and then he
3	5	do n't know	4	i do n't	6	i 'm catching	1	about trudy hey	3	and the pinocchio
4	5	the little people	4	there 's a	6	i do n't	1	and but that	3	he woke up
5	4	and so the	3	nobi do n't	6	you ca n't	1	and the playhouse	3	pinocchio said father
6	4	and the gulliver	3	the first day	6	you want to	1	back this one	2	go to school
7	4	do n't have	2	's a little	4	catching up i	1	better you can	2	he put the
8	4	i have to	2	a little road	4	did n't know	1	but hard cause	2	he went out
9	4	n't have to	2	a sports car	4	do n't need	1	but that thing	2	him and then
10	4	so the gulliver	2	and there is	4	do you want	1	can share need	2	his legs on

付録 H. J の trigram 使用上位 10 位

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	9	no no no	7	do n't	14	i do n't	33	little black sambo	33	and there 's	13	there 's a
2	4	do n't know	7	i do	11	do n't know	13	and little black	24	there 's a	11	i do n't
3	3	i do n't	7	n't know	4	no i do	13	do n't know	19	this is the	10	and there 's
4	3	bag map balu	6	can you	4	you ca n't	11	and the tiger	15	and this is	9	do n't know
5	2	balu n tenis	6	no yes	3	n't know the	11	i do n't	12	and he 's	7	do n't have
6	2	bas taki helikopt	6	speak english	3	nannte iu no	10	you do n't	12	i do n't	6	's lots of
7	2	black and white	6	yes yes	3	no you ca	8	black sambo said	12	over here and	6	now sachiko is
8	2	brown and white	5	do you	3	the zoo no	7	'll eat you	11	and he have	6	there 's lots
9	2	dg bat miluk	5	i have	3	this is i	7	i 'll eat	10	like this and	6	this is a
10	2	dres bag map	5	is coming	3	this is my	7	the jungle and	7	and then he	6	you do n't
月数	21		24		28		31		34			
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖		
1	8	i do n't	11	i do n't	13	let 's pretend	3	do n't have	11	would you like		
2	7	do n't have	7	do n't know	13	this is the	3	gon na be	10	like to say		
3	6	the little people	5	you have to	8	i do n't	3	is gon na	10	you like to		
4	6	the other country	3	here 's your	7	's gon na	3	let 's pretend	9	to say something		
5	6	this is the	3	n't have a	7	i have to	3	they have to	6	the old man		
6	5	do n't know	2	's your car	6	'll take it	2	'm sure glad	5	little girl what		
7	5	went like this	2	's your money	6	i 'll take	2	does n't fit	5	you want some		
8	4	's gon na	2	all the time	6	pretend they ate	2	fit in there	4	time to go		
9	4	has to be	2	and the fire	6	there was a	2	give me the	4	to go home		
10	4	have a war	2	and you have	6	you have to	2	have to move	3	's your name		

付録 I. S の trigram 使用上位 10 位

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	8	no no no	8	ma ma ma	21	i do n't	16	i do n't	44	there 's a	8	came in and
2	3	yes no yes	6	no no no	21	you ca n't	15	we 're going	43	and there 's	8	straw into gold
3	2	berj gud mami	6	this is for	17	you have to	14	're gon na	13	i do n't	8	the little man
4	2	du zat ap	4	and round and	15	do n't know	11	's gon na	13	there 's the	7	i do n't
5	2	faib siks sebun	4	i do n't	11	i ca n't	11	do n't know	13	this is the	7	in and said
6	2	guwd berj gud	4	i play with	10	pretend this is	11	you have to	12	do n't know	6	and i think
7	2	hau du ju	4	is for you	9	ca n't see	10	and he said	12	is the way	6	and that 's
8	2	jo welkm guwd	4	play with ji-yon	9	this is the	10	gon na be	12	the way we	6	and then the
9	2	mai neim iz	4	round and round	9	we have to	10	i ca n't	11	and then she	6	ca n't see
10	2	no mrs vernois	4	you do n't	8	do like this	10	i have to	9	and the father	6	man came in
月数	21		24		28		31		34			
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖		
1	6	i ca n't	10	i do n't	8	i 'll be	4	and this is	7	my name is		
2	6	i do n't	7	do n't know	6	go to the	4	is where they	6	reminds me of		
3	5	go to the	6	're gon na	6	have to go	4	this is where	5	and then he		
4	5	i did n't	5	in the mirror	6	she has to	3	and when they	5	it reminds me		
5	5	you have to	5	let 's see	6	this is the	3	here 's the	4	go to the		
6	4	do n't know	5	you 're gon	5	she did n't	3	see this is	4	the little puppet		
7	4	she has to	4	's see what	5	you have to	3	they have to	4	to go to		
8	3	's gon na	4	a little bit	5	you want to	3	you wan na	3	and then after		
9	3	a little bit	4	and there 's	4	'll be outside	2	'll ride on	3	began to move		
10	3	and he said	4	gon na get	4	'll take the	2	's more kids	3	i 'd better		

付録 J. N の 4-gram 使用上位 10 位

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	5	no no no no	9	i do n't know	4	do you like what	10	i do n't know	4	on the wall there	5	and the wolf came
2	3	no yes no yes	3	go to school tommorrow	3	go to school tommorrow	7	gon na eat you	3	and he put the	4	's house and the
3	2	ai maus eapo t	3	no i do n't	3	i do n't know	6	eat you tiger said	3	and one boy is	4	sheep 's house and
4	2	blak sri ho haib	3	yes i do n't	3	what do you like	6	is pulling the old	3	and the dog is	4	the sheep 's house
5	2	di monki taiga laion	2	's boy scout scarf	2	's blanket me nobuki	6	na eat you tiger	3	and the father is	3	and this day is
6	2	elehant zi bla hipopotms	2	and two yes two	2	's brothers and sisters	5	little black sambo i	3	boy and the girl	3	came out from the
7	2	hamm naih ho k	2	boy scout scarf no	2	a small tree and	4	a little bit more	3	has brown hair and	3	came to the sheep
8	2	hea ai maus eapo	2	do n't know three	2	and watch it everyday	4	black sambo i gon	3	i do n't know	3	do the same thing
9	2	hed hea ai maus	2	do n't know yes	2	beautiful day no spring	4	do n't eat me	3	is sit on the	3	pig 's house and
10	2	hi tip hi tip	2	do you like this	2	become a rich man	4	eat me i give	3	is sitting on the	3	said come out or

続き

月数	21		24		28		31		34	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	5	i do n't know	4	and there 's a	6	i 'm catching up	1	's your name hi	2	and the pinochio said
2	4	to the other island	4	i do n't know	5	i do n't know	1	about trudy hey turtle	2	go to school and
3	3	do n't have to	2	does n't have a	4	'm catching up i	1	and but that thing	2	his legs on the
4	2	all the ships that	2	if you get you	4	do you want to	1	and the playhouse here	2	put his legs on
5	2	and so the king	2	is a sports car	4	i did n't know	1	back this one is	1	'll be a real
6	2	and that 's my	2	left and there 's	4	no you ca n't	1	better you can share	1	'll buy you a
7	2	and the other guys	2	no i do n't	3	do n't know but	1	but hard cause what	1	'm just trying to
8	2	came from the other	2	she does n't have	2	' in the bedroom	1	but that thing does	1	's head and everybody
9	2	car right over here	2	there 's a little	2	' what ' in	1	can share need to	1	's heavy so he
10	2	do n't know what	2	this car is a	2	'll get some i	1	cause what now this	1	's is making the

付録 K. J の 4-gram 使用上位 10 位

月数	3	6	9	12	15	18
順位	頻度 連鎖	頻度 連鎖	頻度 連鎖	頻度 連鎖	頻度 連鎖	頻度 連鎖
1	5 no no no no	7 i do n't know	11 i do n't know	9 and little black sambo	19 and there 's a	8 i do n't know
2	3 i do n't know	3 four five six seven	4 no i do n't	9 i do n't know	6 and he have a	6 there 's lots of
3	2 bag map balu n	3 i play game with	3 do n't know the	8 little black sambo said	6 and this is the	4 let 's make a
4	2 balu n tennis blakbo	3 one two three four	3 no you ca n't	7 i 'll eat you	6 i think this is	3 and there 's a
5	2 bas taki helikopt eskale	3 play game with rollo	3 tte nannte iu no	5 and the doll is	5 and there 's some	3 and there 's lots
6	2 dres bag map balu	3 three four five six	2 's name is sadako	5 and the tiger said	5 he 's gon na	3 did n't come out
7	2 grey brown and white	3 two three four five	2 and mrs mcMahon and	5 black sambo i 'll	4 bad guys ' fort	3 do n't get your
8	2 hat dg bat milik	2 and last name is	2 boku wa nichiyoga	5 do n't eat me	4 guys ' fort and	3 do n't know i
9	2 ho ku spu n	2 and me speak english	2 daisukidesu i like sunday	5 i 'll give you	4 i do n't know	3 get out of here
10	2 map balu n tennis	2 and this puzzle and	2 day do you like	5 little black sambo i	4 once upon a time	3 get your head on

続き

月数	21		24		28		31		34	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	5	i do n't know	7	i do n't know	6	i 'll take it	3	is gon na be	10	would you like to
2	4	do n't have a	2	and you have to	5	i do n't know	2	does n't fit in	10	you like to say
3	4	the other country and	2	are supposed to go	5	let 's pretend they	2	give me the boat	9	like to say something
4	4	you went like this	2	go like wait i	4	'll take it away	2	here this is the	4	time to go home
5	3	the barber and the	2	got out of the	4	's pretend they ate	2	house is gon na	3	girl what 's your
6	3	the king said to	2	have to keep that	4	i know i did	2	i 'm sure glad	3	hi little girl what
7	3	the land where all	2	he got out of	4	pretend they ate them	2	if she were here	3	little girl what 's
8	3	to the land where	2	here 's your money	4	this is the hotel	2	n't fit in there	3	what 's your name
9	3	who 's gon na	2	i do n't have	4	up and up and	1	'cause i wan na	2	'm the cereal man
10	2	's raining outside it	2	like wait i do	3	and then when the	1	'll be lonely i	2	's your name oh

付録 L. S の 4-gram 使用上位 10 位

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	4	no no no no	7	ma ma ma ma	11	i do n't know	8	i do n't know	25	and there 's a	7	came in and said
2	2	du zat ap do	4	i play with ji-yon	6	do n't have to	7	and there 's a	12	i do n't know	5	man came in and
3	2	faib siks sebun eit	4	round and round and	5	i ca n't see	7	we 're going to	12	is the way we	5	then the little man
4	2	guwd berj gud mami	4	this is for you	4	ca n't see it	7	we 're gon na	12	this is the way	4	i 'll give you
5	2	jo welkm guwd berj	3	gon na be terry	4	no you ca n't	6	i 'm gon na	6	and there 's the	4	into gold by morning
6	2	plijz jo welkm guwd	3	i 'm gon na	4	you ca n't see	5	she 's gon na	5	on the right hand	4	said i 'll give
7	2	welkm guwd berj gud	3	no no no no	4	you do n't have	5	you 're gon na	5	there 's a bush	4	the little man said
8	1	a hous do n	3	o k o k	4	you do n't know	4	'll eat you up	4	and there 's still	4	then the king said
9	1	a little kam hir	2	and this is for	4	you have to do	4	do it like this	4	corner there 's a	4	then the next day
10	1	a no yes no	2	do n't kick do	3	and this is the	4	he 's gon na	4	do you call this	4	this straw into gold

月数	21		24		28		31		34	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	4	i do n't know	7	i do n't know	4	go to the bathroom	4	this is where they	5	it reminds me of
2	3	did n't go like	5	you 're gon na	4	i 'll be outside	2	'll ride on you	3	began to move and
3	3	i did n't go	4	're gon na get	4	i 'll take the	2	and this is where	3	me of wild hickery
4	3	to go to the	4	and there 's a	4	i do n't know	2	does n't come off	3	of wild hickery nuts
5	2	's gon na be	4	let 's see what	4	put them over there	2	i 'll ride on	3	reminds me of wild
6	2	and pretgel said that	3	at herself in the	4	to go to the	2	is too big for	3	would you like to
7	2	did just one apple	3	herself in the mirror	4	you have to go	2	it 's my birthday	2	and the girl said
8	2	do n't have to	3	now let 's see	4	you want to go	2	see this is where	2	and then after he
9	2	go to the circus	2	'm very sorry that	3	's going to be	2	that does n't come	2	better go to the
10	2	have to make that	2	're saying that to	3	did n't even know	2	there 's more kids	2	do n't have enough

付録 M. N の品詞 bigram 上位 10 位の推移

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	141	nn1 nn1	78	jp jp	183	jp jp	95	at nn1	177	at nn1	102	at nn1
2	44	uh uh	36	uh uh	32	ppis1 vv0	89	jj nn1	165	nn1 cc	73	nn1 cc
3	41	mc mc	27	ppis1 vv0	31	at nn1	77	nn1 nn1	64	ii at	37	cc at
4	25	jp jp	21	nn1 ppis1	30	jj nn1	64	nn1 cc	64	nn2 cc	35	nn1 vvd
5	13	nn1 vv0	15	xx vvi	28	appge nn1	64	nn1 vbz	60	cc pphs1	31	ii at
6	10	np1 np1	14	np1 np1	24	uh uh	42	ii at	58	nn1 vbz	30	to vvi
7	10	vv0 vv0	14	vv0 nn1	21	nn1 ppis1	42	nn1 vvd	57	jj nn1	26	jj nn1
8	9	jj jj	13	nn1 nn1	19	dd1 vbz	38	xx vvi	47	cc at	26	pphs1 vvd
9	9	uh jj	13	nn1 uh	19	nn1 jp	37	appge nn1	46	pphs1 vvd	25	appge nn1
10	9	uh nn1	13	ppis1 vd0	19	nn1 uh	36	cc nn1	45	cc nn1	24	dd1 vbz
月数	21		24		28		31		34			
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖		
1	79	at nn1	23	at nn1	37	at nn1	2	dd1 pn1	19	pphs1 vvd		
2	48	ii at	14	xx vvi	22	ii at	2	pn1 vbz	17	nn1 cc		
3	42	nn1 cc	13	vd0 xx	21	to vvi	2	vdz xx	15	cc rt		
4	42	to vvi	12	jj nn1	21	xx vvi	2	vm vvi	14	at nn1		
5	35	at1 nn1	12	to vvi	18	ppis1 vv0	2	vvi cc	10	to vvi		
6	33	nn1 nn1	10	ii at	16	nn1 ii	2	vvi rp	9	nn1 vvd		
7	26	appge nn1	10	nn1 nn1	16	uh ppis1	2	xx vvi	8	cc at		
8	26	at jj	10	vbz at1	16	vm vvi	1	appge nn1	8	cc pphs1		
9	26	jj nn1	9	ex vbz	15	ppis1 vvd	1	at nn1	7	vvd rp		
10	23	jp jp	9	nn1 uh	13	nn1 uh	1	cc at	6	np1 vvd		

付録 N. J の品詞 bigram 上位 10 位の推移

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	143	nn1 nn1	28	nn1 nn1	91	jp jp	128	at nn1	264	at nn1	121	at nn1
2	51	jp jp	25	mc mc	37	nn1 cc	122	nn1 cc	261	nn1 cc	70	nn1 cc
3	33	uh uh	20	jp jp	32	nn1 uh	97	jj nn1	109	at1 nn1	58	at1 nn1
4	17	mc mc	19	vv0 nn1	30	dd1 vbz	59	nn1 vvd	105	jj nn1	46	xx vvi
5	16	jj jj	18	uh uh	29	nn1 nn1	43	cc at	103	ii at	44	ex vbz
6	12	nn1 jj	12	uh nn1	27	appge nn1	41	appge nn1	88	nn1 nn1	42	nn1 nn1
7	12	nn1 vv0	11	nn1 cc	26	at nn1	41	ii at	87	appge nn1	41	ii at
8	10	vv0 vv0	9	nn1 uh	26	nn1 vbz	39	xx vvi	81	dd1 vbz	41	to vvi
9	9	jj nn1	8	ppy vvi	24	uh uh	35	nn1 uh	77	cc pphs1	37	appge nn1
10	9	np1 np1	7	mc uh	19	jj nn1	34	dd1 vbz	73	cc at	36	dd1 vbz
月数	21		24		28		31		34			
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖		
1	131	at nn1	41	at nn1	160	at nn1	23	at nn1	26	to vvi		
2	55	nn1 cc	19	xx vvi	71	to vvi	18	to vvi	22	appge nn1		
3	50	ii at	18	nn1 nn1	58	ii at	11	vm vvi	22	jj nn1		
4	42	at1 nn1	16	vd0 xx	54	nn1 cc	10	dd1 vbz	18	at nn1		
5	35	to vvi	15	appge nn1	50	xx vvi	8	pph1 vbz	18	nn1 nn1		
6	32	jj nn1	15	to vvi	41	nn1 nn1	7	at nn2	18	pphs1 vvd		
7	31	cc at	12	ii at	40	vm vvi	7	dd1 nn1	18	ppy vvi		
8	31	ppis1 vv0	11	ppis1 vd0	38	at1 nn1	7	ppis1 vv0	15	vm ppy		
9	23	nn1 nn1	10	uh uh	38	ppis1 vv0	7	vbz at	13	at1 nn1		
10	19	at jj	9	nn1 ii	37	dd1 vbz	7	vvgk to	13	nn1 cc		

付録 O. S の品詞 bigram 上位 10 位の推移

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	80	nn1 nn1	54	uh uh	107	jp jp	109	nn1 cc	305	at nn1	96	at nn1
2	28	uh uh	30	nn1 nn1	73	dd1 vbz	107	at nn1	190	nn1 cc	60	to vvi
3	17	mc mc	23	dd1 vbz	64	vd0 xx	85	to vvi	129	at1 nn1	55	nn1 cc
4	11	nn1 uh	19	nn1 uh	64	xx vvi	66	at1 nn1	126	ii at	52	appge nn1
5	11	vv0 vv0	19	nn1 vv0	55	appge nn1	62	ii dd1	87	ex vbz	51	jj nn1
6	10	uh nn1	18	ppis1 vv0	55	uh uh	59	xx vvi	82	vbz vvg	49	at1 nn1
7	9	jj jj	18	uh nn1	54	ppis1 vv0	54	appge nn1	74	cc at	45	ppis1 vv0
8	9	vv0 nn1	18	vd0 xx	54	to vvi	54	dd1 vbz	71	nn1 vbz	43	pphis1 vvd
9	8	nn1 jj	16	nn1 rr	53	ppy vm	53	ppis1 vv0	62	at nn2	41	ii at
10	8	nn1 vv0	15	vv0 nn1	43	vm xx	51	ii at	62	nn1 nn1	39	nn1 vvd
月数	21		24		28		31		34			
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖		
1	87	at nn1	43	at nn1	80	at nn1	25	at nn1	38	at nn1		
2	42	nn1 cc	38	at1 nn1	68	to vvi	24	to vvi	30	pphis1 vvd		
3	40	at1 nn1	34	to vvi	39	ppis1 vm	18	ppis1 vv0	26	jj nn1		
4	40	ii at	31	uh uh	36	ii at	18	vm vvi	26	to vvi		
5	33	to vvi	25	xx vvi	30	uh uh	17	dd1 vbz	17	nn1 cc		
6	32	ppis1 vv0	23	ii at	30	vm vvi	17	pphis2 vv0	15	cs pphis1		
7	31	jj nn1	21	ppis1 vv0	29	ppis1 vv0	15	ppy vv0	13	cc rt		
8	30	np1 vvz	20	nn1 uh	26	xx vvi	13	appge nn1	13	ii at		
9	27	pphis1 vvd	18	nn1 cc	22	nn1 cc	13	xx vvi	13	vm vvi		
10	27	xx vvi	18	vd0 xx	22	pphis1 vbz	10	at1 nn1	12	appge nn1		

付録 P. N の品詞 trigram 使用上位 10 位

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	113	nn1 nn1 nn1	56	jp jp jp	137	jp jp jp	35	ii at nn1	56	ii at nn1	25	cc at nn1
2	36	mc mc mc	17	uh uh uh	18	nn1 jp jp	27	vd0 xx vvi	38	cc at nn1	22	nn1 cc at
3	27	uh uh uh	13	ppis1 vd0 xx	14	nn1 ppis1 vv0	26	mc mc mc	37	at nn1 cc	20	at nn1 cc
4	18	jp jp jp	13	vd0 xx vvi	13	vd0 ppy vvi	25	nn1 vbz vvg	33	jj nn1 cc	20	at nn1 vvd
5	7	nn1 nn1 vv0	12	nn1 ppis1 vv0	11	jp jp nn1	21	at nn1 nn1	33	nn1 cc pphs1	18	ii at nn1
6	6	np1 np1 np1	8	vd0 ppy vvi	10	ii at nn1	19	at jj nn1	30	cc pphs1 vvd	13	ddl vbz at
7	5	nn1 nn1 nn2	7	np1 np1 np1	10	nn2 nn2 nn2	19	at nn1 cc	28	nn1 cc at	12	ge nn1 cc
8	5	np1 nn1 nn1	7	ppis1 vv0 nn1	9	at1 jj nn1	18	nn1 ge nn1	28	nn1 vbz vvg	11	cc pphs1 vvd
9	5	npd1 npd1 npd1	6	nn1 jp jp	9	np1 cc np1	18	rr jj nn1	27	at nn1 vbz	11	vbz at nn1
10	5	vv0 uh uh	6	nn1 uh uh	8	ppis1 vd0 xx	15	nn1 cc nn1	24	nn1 cc nn1	11	vvd ii at
月数	21		24		28		31		34			
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖		
1	28	ii at nn1	11	vd0 xx vvi	14	ii at nn1	2	ddl pn1 vbz	7	at nn1 vvd		
2	16	jp jp jp	6	at1 jj nn1	12	vd0 xx vvi	2	vdz xx vvi	6	cc pphs1 vvd		
3	14	at jj nn1	6	cc ex vbz	11	nn1 ii at	1	appge nn1 uh	6	pphs1 vvd rp		
4	11	ii at jj	6	ii at nn1	8	at nn1 rt	1	at nn1 rl	5	cc rt at		
5	10	vvd to vvi	5	ex vbz at1	8	ppis1 vbm vvg	1	cc at nn1	5	vvd to vvi		
6	9	at nn1 nn1	4	cc ppy vv0	8	ppis1 vm vvi	1	cc ccb dd1	4	cc rt pphs1		
7	9	at nn1 vvd	4	ppis1 vd0 xx	8	vh0 to vvi	1	cc vvi rp	4	nn1 cc pphs1		
8	9	at np1 vvd	4	xx vvi rp	7	uh ppis1 vv0	1	ccb dd1 nn1	4	nn1 cc rt		
9	9	nn1 ii at	3	at md nnt1	6	at nn1 uh	1	ccb rr vv0	3	at nn1 cc		
10	9	nn1 vbz vvg	3	at nn1 nn1	6	ppis1 vd0 xx	1	dd io dd1	3	at np1 vvd		

付録 Q. J の品詞 trigram 使用上位 10 位

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	116	nn1 nn1 nn1	21	mc mc mc	64	jp jp jp	32	cc at nn1	83	ii at nn1	31	ii at nn1
2	43	jp jp jp	10	jp jp jp	15	ppis1 vd0 xx	32	ii at nn1	80	at nn1 cc	26	vd0 xx vvi
3	19	uh uh uh	8	nn1 nn1 cc	13	nn1 cc nn1	29	at nn1 cc	46	nn1 cc at	19	at nn1 cc
4	16	mc mc mc	7	ppis1 vd0 xx	10	ii at nn1	29	rr jj nn1	42	cc at nn1	19	vbz at1 nn1
5	11	jj jj jj	7	ppy vvi nn1	10	vd0 xx vvi	28	jj nn1 cc	40	nn1 cc pphs1	14	at md nn1
6	9	nn1 nn1 jj	7	uh uh uh	9	nn1 nn1 uh	26	nn1 cc at	40	vbz at1 nn1	14	nn1 cc at
7	7	nn1 nn1 vv0	7	vd0 xx vvi	8	nn1 vbz np1	25	vd0 xx vvi	39	cc ex vbz	13	at1 nn1 nn1
8	6	uh uh nn1	6	nn1 cc nn1	7	cc nn1 cc	19	jj nn1 vvd	38	at1 nn1 cc	13	ex vbz at1
9	5	jj nn1 nn1	6	vm ppy vvi	6	appge nn1 vbz	18	at nn1 vvd	32	pphs1 vbz vvg	13	vbz at nn1
10	5	nn1 vv0 vv0	5	nn1 vbz np1	6	at nn1 nn1	18	nn1 ge nn1	30	nn1 cc nn1	12	at1 jj nn1
月数	21		24		28		31		34			
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖		
1	40	ii at nn1	12	vd0 xx vvi	43	ii at nn1	6	vm21 vm22 vvi	12	vm ppy vvi		
2	23	at nn1 cc	11	ppis1 vd0 xx	24	vd0 xx vvi	4	to vvi dd1	11	vvi to vvi		
3	22	cc at nn1	10	ii at nn1	23	ppis1 vm vvi	4	vbz at nn1	10	ppy vvi to		
4	18	nn1 cc at	6	at nn1 nn1	23	vh0 to vvi	4	vh0 to vvi	9	to vvi pn1		
5	12	at nn1 cs	6	at nn1 pphs1	22	at nn1 cc	4	vvgk to vvi	7	at1 jj nn1		
6	12	vbz at nn1	6	nn1 nn1 nn1	22	vbz at nn1	3	appge nn1 vbz	6	appge nn1 uh		
7	11	vvi at nn1	6	vbz appge nn1	21	vm21 vm22 vvi	3	at jj nn1	6	at jj nn1		
8	9	at jj nn1	5	at nn1 vvd	19	ppis1 vd0 xx	3	at nn1 to	6	cc pphs1 vvd		
9	9	at1 jj nn1	5	at1 jj nn1	17	dd1 vbz at	3	dd1 vbz at	6	ii appge nn1		
10	9	dd1 vbz at	5	nn1 at nn1	15	at1 jj nn1	3	ii at nn1	6	vbz appge nn1		

付録 R. S の品詞 trigram 使用上位 10 位

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	60	nn1 nn1 nn1	31	uh uh uh	80	jp jp jp	39	mc mc mc	98	ii at nn1	29	ii at nn1
2	17	uh uh uh	12	vd0 xx vvi	30	vm xx vvi	32	ii at nn1	76	at nn1 cc	22	at nn1 vvd
3	14	mc mc mc	10	nn1 uh uh	29	vd0 xx vvi	30	vvgk to vvi	51	cc at nn1	18	at jj nn1
4	6	jj nn1 nn1	8	jp jp jp	23	ppis1 vd0 xx	28	vd0 xx vvi	49	ex vbz at1	14	at nn1 cc
5	4	nn1 nn1 jj	8	uh uh nn1	23	vh0 to vvi	25	at nn1 cc	44	at1 nn1 cc	14	jj nn1 vvd
6	4	nn1 nn1 nn2	6	ddl vbz if	22	ppy vm xx	22	vh0 to vvi	44	cc ex vbz	14	vd0 xx vvi
7	4	nn1 nn1 uh	6	nn1 nn1 nn1	18	np1 cc np1	21	ppis1 vm vvi	41	vbz at1 nn1	13	at1 jj nn1
8	4	nn1 uh nn1	6	nn1 nn1 rr	18	ppy vh0 to	21	vbz at1 nn1	40	nn1 vbz vvg	13	nn1 vvd rp
9	4	nn1 vv0 vv0	5	ddl vbz at	17	ii at nn1	21	vm xx vvi	37	at nn1 vbz	12	rt at nn1
10	4	nn2 nn1 nn1	4	cc jj cc	17	nn1 cc nn1	19	vbr vvgk to	36	nn1 cc at	12	rt pphs1 vvd
月数	21		24		28		31		34			
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖		
1	34	ii at nn1	20	ii at nn1	20	ii at nn1	8	vv0 to vvi	12	cs pphs1 vvd		
2	16	np1 vvz mc	13	vd0 xx vvi	14	ppis1 vm vvi	6	cs pphs2 vv0	11	ii at nn1		
3	14	at1 jj nn1	13	vvgk to vvi	13	vbz at nn1	6	ii at nn1	10	vvd to vvi		
4	13	mc np1 vvz	12	uh uh uh	10	vh0 to vvi	6	vbz at nn1	9	at jj nn1		
5	13	vvz mc np1	10	ppis1 vd0 xx	10	vm21 vm22 vvi	5	at nn1 ppis1	9	at nn1 cc		
6	12	ppis1 vm vvi	8	at1 jj nn1	10	vvgk to vvi	5	nn1 ppy vv0	9	nn1 vbz np1		
7	10	cc pphs1 vvd	8	vbr vvgk to	8	at jj nn1	5	pphs2 vv0 rp	8	appge nn1 vbz		
8	9	at nn1 cc	7	at nn1 nn1	8	at jj pn1	5	ppis1 vm vvi	8	pphs1 vvd at		
9	9	at nn1 rr	7	vbz at1 nn1	8	at nn1 cc	5	vh0 to vvi	8	ppis1 vm vvi		
10	9	nn1 ii at	7	vdd xx vvi	8	at nn1 rr	5	vm xx vvi	7	at1 jj nn1		

付録 S. N の品詞 4-gram 使用上位 10 位

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	98	nn1 nn1 nn1 nn1	39	jp jp jp jp	106	jp jp jp jp	20	mc mc mc mc	26	nn1 cc at nn1	13	nn1 cc at nn1
2	32	mc mc mc mc	13	ppis1 vd0 xx vvi	12	nn1 jp jp jp	14	nn1 vbz vvg at	17	at nn1 ex vbz	12	cc at nn1 vvd
3	16	uh uh uh uh	6	uh ppis1 vd0 xx	9	jp jp jp nn1	11	ii at nn1 nn1	15	at nn1 vbz vvg	9	dd1 vbz at nn1
4	12	jp jp jp jp	5	jp jp jp nn1	6	appge nn1 ge nn1	10	at nn1 nn1 vbz	15	ii at nn1 cc	8	at nn1 cc at
5	4	nn1 nn1 nn1 nn2	5	uh uh ppis1 vd0	6	nn1 ppis1 vv0 nn1	10	ppis1 vd0 xx vvi	15	ii at nn1 ex	8	ii at nn1 cc
6	4	nn1 nn1 nn1 vv0	5	uh uh uh uh	5	appge nn1 ppis1 vv0	8	jj nn1 jj nn1	15	nn1 ii at nn1	7	nn1 ii at nn1
7	4	nn1 vv0 uh uh	4	nn1 jp jp jp	5	dd1 vbz appge nn1	8	vbz vvg at nn1	14	cc at nn1 vbz	6	at nn1 vbz vvg
8	4	np1 np1 np1 np1	4	nn1 ppis1 vv0 nn1	5	jp jp nn1 jp	8	vvg at jj nn1	14	nn1 cc pphs1 vvd	6	nn1 ge nn1 cc
9	3	nn1 nn1 jj jj	4	nn1 uh uh uh	5	jp nn1 jp jp	7	at nn1 cc nn1	13	cc jj nn1 cc	5	at nn1 vvd ii
10	3	nn1 nn1 nn2 nn1	4	nn1 vd0 ppy vvi	5	nn2 nn2 nn2 nn2	7	at nn1 ge nn1	13	jj nn1 cc jj	5	at nn1 vvd rp
月数	21		24		28		31		34			
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖		
1	11	jp jp jp jp	5	cc ex vbz at1	7	nn1 ii at nn1	1	appge nn1 uh ppis1	3	cc at nn1 vvd		
2	8	ii at jj nn1	4	ppis1 vd0 xx vvi	6	ppis1 vbm vvg rp	1	at nn1 rl rrq	3	cc at np1 vvd		
3	8	nn1 ii at nn1	4	vd0 xx vvi rp	6	ppis1 vd0 xx vvi	1	cc at nn1 rl	3	nn1 cc pphs1 vvd		
4	6	vvg ii at nn1	3	ex vbz at1 jj	4	at nn1 rt ppis1	1	cc ccb dd1 nn1	3	pphs1 vvd to vvi		
5	5	ex vbz at1 nn1	3	vbz at1 jj nn1	4	ii at nn1 rt	1	cc vvi rp dd1	2	appge nn2 ii at		
6	5	pphs1 vvd to vvi	3	vbz at1 nn2 nn1	4	ppis1 vdd xx vvi	1	ccb dd1 nn1 vdz	2	cc pphs1 vvd ii		
7	5	ppis1 vd0 xx vvi	2	at nn1 uh uh	4	ppis2 vd0 xx vvi	1	ccb rr vv0 ddq	2	cc rt at nn1		
8	4	ii at nn1 cc	2	at1 jj nn1 ii	4	ppis2 vh0 to vvi	1	dd io dd1 vv0	2	cc rt pphs1 vvd		
9	4	ii at nn1 mc	2	cc ppy vv0 at	4	ppy vm xx uh	1	dd pphs1 vvd dd	2	cs at nn1 vvd		
10	4	ii at nn1 nn1	2	dd1 nn1 vbz at1	4	ppy vvi to vvi	1	dd1 dd1 vdz xx	2	nn1 cc rt pphs1		

付録 T. J の品詞 4-gram 使用上位 10 位

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	100	nn1 nn1 nn1 nn1	17	mc mc mc mc	42	jp jp jp jp	22	nn1 cc at nn1	31	ii at nn1 cc	9	ex vbz at1 nn1
2	35	jp jp jp jp	7	jp jp jp jp	10	ppis1 vd0 xx vvi	13	ii at nn1 cc	29	nn1 cc at nn1	8	nn1 cc ex vbz
3	15	mc mc mc mc	7	ppis1 vd0 xx vvi	5	nn1 jp jp jp	13	rr jj nn1 vvd	27	ex vbz at1 nn1	8	ppis1 vd0 xx vvi
4	11	uh uh uh uh	6	nn1 nn1 cc nn1	5	uh ppis1 vd0 xx	12	cc at nn1 vvd	22	cc ex vbz at1	7	cc at md nn1
5	9	jj jj jj jj	5	vm ppy vvi nn1	4	cc nn1 cc nn1	12	cc rr jj nn1	19	at nn1 cc at	7	ii at nn1 cc
6	5	nn1 nn1 nn1 jj	4	md nn1 vbz np1	4	ddl vbz appge nn1	12	ppis1 vm vvi ppy	18	nn1 cc ex vbz	7	nn1 cc at nn1
7	5	nn1 nn1 vv0 vv0	4	nn1 cc nn1 nn1	4	md nn1 vbz np1	11	at nn1 cc at	17	vvg ii at nn1	7	vbz at1 nn1 nn1
8	4	nn1 jj cc jj	4	nn1 vm ppy vvi	4	nn1 cc nn1 cc	10	jj nn1 cc jj	16	nn1 ii at nn1	6	at nn1 cc at
9	4	nn1 nn1 jj jj	3	cc nn1 nn1 cc	4	nn1 ppis1 vd0 xx	10	ppis1 vd0 xx vvi	15	cc pphs1 vbz vvg	6	ex vbz pn io
10	3	jj nn1 nn1 nn1	3	mc1 mc mc mc	4	np1 cc nnb np1	9	nn1 cc rr jj	15	nn1 cc pphs1 vvd	6	nn1 cc at1 nn1
月数	21		24		28		31		34			
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖		
1	13	nn1 cc at nn1	7	ppis1 vd0 xx vvi	15	ppis1 vd0 xx vvi	3	at nn1 to vvi	10	ppy vvi to vvi		
2	12	at nn1 cc at	4	nn1 nn1 nn1 nn1	13	ddl vbz at nn1	3	nn1 vbz vvgk to	10	vm ppy vvi to		
3	8	ddl vbz at nn1	4	ppy vh0 to vvi	12	ii at nn1 cc	3	vbz vvgk to vbi	9	vvi to vvi pn1		
4	8	ii at nn1 cc	3	ii at nn1 pphs1	10	ppis1 vm vvi pph1	2	appge nn1 vbz vvgk	5	ddq vbz appge nn1		
5	7	ii at nn1 cs	3	rl vbz appge nn1	8	ppy vh0 to vvi	2	at nn1 nn1 vbz	5	nn1 ddq vbz appge		
6	7	to vvi at nn1	3	rrq vbz appge nn1	8	vh0 to vvi at	2	at nn1 vm21 vm22	5	vbz appge nn1 uh		
7	6	ii at jj nn1	2	at nn1 ge nn1	8	vvi ii at nn1	2	at nn2 rr vv0	5	vbz at1 jj nn1		
8	6	nn1 ii at nn1	2	at nn1 nn1 at	7	to vvi at nn1	2	cs pphs1 vbdr rl	4	at jj nn1 vvd		
9	5	at jj nn1 cc	2	at nn1 nn1 cc	7	vvd ii at nn1	2	cs pphs2 vm vvi	4	jj nn1 ddq vbz		
10	5	ppis1 vd0 xx vvi	2	at nn1 pphs1 vvd	6	at nn1 cc rt	2	cs ppis1 vm vvi	4	mc nnu nn1 nn1		

付録 U. S の品詞 4-gram 使用上位 10 位

月数	3		6		9		12		15		18	
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖
1	47	nn1 nn1 nn1 nn1	20	uh uh uh uh	62	jp jp jp jp	32	mc mc mc mc	47	ii at nn1 cc	12	at jj nn1 vvd
2	12	mc mc mc mc	5	jp jp jp jp	13	ppis1 vd0 xx vvi	11	ppis1 vd0 xx vvi	39	ex vbz at1 nn1	9	ii at nn1 cc
3	10	uh uh uh uh	4	dd1 vbz at nn1	13	ppy vm xx vvi	11	ppis2 vbr vvgk to	29	nn1 cc at nn1	9	rt at nn1 vvd
4	5	jj nn1 nn1 nn1	4	dd1 vbz if ppy	10	ppis1 vm xx vvi	11	vbr vvgk to vvi	27	cc ex vbz at1	8	nn1 vvd rp cc
5	4	nn1 nn1 nn1 jj	4	nn1 uh uh uh	10	ppy vh0 to vvi	10	np1 cc np1 cc	26	nn1 ii at nn1	7	jj nn1 vvd rp
6	3	nn1 nn1 nn1 nn2	4	ppis1 vd0 xx vvi	9	dd1 vbz at1 nn1	10	ppy vh0 to vvi	25	cc at nn1 vbz	7	ppy vm xx vvi
7	3	nn1 nn1 nn2 nn1	4	ppis1 vv0 iw np1	9	np1 cc np1 cc	10	vbz vvgk to vvi	24	at nn1 cc at	7	vvd rp cc vvd
8	3	nn1 nn2 nn1 nn1	4	uh uh uh nn1	8	nn1 ppis1 vd0 xx	9	ii at nn1 cc	19	at nn1 vbz vvg	6	at nn1 vvd rp
9	3	nn2 nn1 nn1 nn1	3	dd1 vbz appge nn1	7	cc np1 cc np1	9	pphis1 vbz vvgk to	19	dd1 vbz at nn1	6	rt at jj nn1
10	3	uh uh uh nn1	3	jj cc jj cc	7	dd1 vbz at nn1	8	cc np1 cc np1	19	vvg ii at nn1	6	rt at md nnt1
月数	21		24		28		31		34			
順位	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖	頻度	連鎖		
1	13	np1 vvz mc np1	8	ppis1 vd0 xx vvi	6	pphis1 vhz to vvi	4	at nn1 ppis1 vm	8	appge nn1 vbz np1		
2	12	vvz mc np1 vvz	5	ex vbz at1 nn1	6	ppis1 vm vvi at	4	dd1 vbz rrq pphis2	5	cc rt pphis1 vvd		
3	10	mc np1 vvz mc	5	ppy vbr vvgk to	6	vbz vvgk to vbi	4	vbz rrq pphis2 vv0	5	pph1 vvz ppio1 io		
4	8	nn1 ii at nn1	5	uh uh uh uh	5	at nn1 vv0 ii	3	cs pphis2 vv0 nn1	5	pphis1 vvd at nn1		
5	7	vvi ii at nn1	5	vbr vvgk to vvi	5	ii at nn1 cc	3	dd1 vbz rg jj	4	cs pphis1 vvd at		
6	6	ppis1 vm xx vvi	4	at1 jj nn1 uh	5	ppis2 vbr vvgk to	3	pphis2 vh0 to vvi	4	nn2 vvd to vvi		
7	5	np1 vvz nn1 np1	4	cc ex vbz at1	5	ppy vh0 to vvi	3	ppis1 vm xx vvi	4	pphis1 vvd to vvi		
8	5	pphis1 vhz to vvi	4	ii ppx1 ii at	5	to vvi ii at	3	ppy vv0 to vvi	4	ppio1 vv0 np1 ppio1		
9	5	ppis1 vd0 xx vvi	4	nn1 cc ex vbz	5	vbr vvgk to vvi	3	rrq vbz at nn1	4	rt cs pphis1 vvd		
10	5	ppis1 vdd xx vvi	4	ppx1 ii at nn1	5	vvi ii at nn1	3	vbz at1 nn1 uh	4	vvd to vvi cc		

Part 3 小池コーパス発話分析作業報告（2011~2013 年度）及び日本人児童の英語発話データに見られる語彙の発達

蒲原 順子 担当

はじめに

このプロジェクトは、日本人児童 3 名の英語の長期発話データをコーパス化し、第二言語習得的観点から分析を行なうことを目的としている。この発話データは、このプロジェクトの発起人である小池生夫氏の博士論文のために 1972 年から約 2 年半に渡って収集されたものである。但し、小池の博士論文では、1 年分のデータのみを扱っている。今回のプロジェクトでは、その後の 1 年半分のデータを含めた全データを対象としてこの 3 名の児童の英語の発達を調査分析する。分析項目は、現在のところ、語彙、品詞の発達、などであるが、今後は統語の発達やコミュニケーションストラテジーと統語の発達における相関などにも対象を広げていく計画である。以下、現在に至るまでの 3 年間の活動を報告する。

3.1.1 2011 年の活動：発話データの基本設定

発話データのコーパス化に当たっては、過去 2004 年から 2005 年に渡って手書きのスクリーンショットで保存されていたオリジナル版を電子化している。(資料 1 参照) ファイル数は 139 で、ファイルによってサイズが異なるが、一つのファイルは一回のインタビューで採集した発話である。

3.1.2 品詞タグと統語タグの付与

2011 年には、基本的な設定として、すでにコーパス化された電子版のテキストデータにあるひとつひとつの単語に品詞のタグを付与した。これは、3 人の子どもの語彙の発達を様々な観点からの統計的な分析を可能にするものである。品詞のタグ付与には、Constituent Likelihood Automatic Word-tagging System (CLAWS) というランカスター大学の University Centre for Computer Corpus Research on Language (UCREL) が開発した品詞のタグ付与ツールを使用した。品詞のタグは、同じく UCREL が開発した CLAWS 7 タグセットを使用した。これは一つの品詞をさらに階層化して名前を付けたもので、全部で 149 個ある (資料 2 参照)。CLAWS 7 タグセットの正式な使用にはライセンスが必要となるが、東京外国語大学の投野由紀夫教授に依頼し、氏の所有するライセンスを使い CLAWS 7 タグセットを用いてテキストデータをタグ付けした。ご多忙にも関わらず、快くタグ付与の便宜を計らって下さった投野氏のご助力に深く感謝の意を表したい。

また、機械で自動的に付けられたタグの中には、品詞情報が間違っているものや、日本語に英語の品詞タグがついているものがあり、目視でのタグの訂正が必要であった。一方、文型についてのタグは機械でつけるものが実在しないため、1 文ずつ見て判断しマニュアル

でタグ付けをすることになった。これら2種類の作業は、投野由紀夫氏のご協力により東京外国語大学の学生14名に作業を依頼した。

3.1.3 発話データのタグ付け作業

まず、作業にかかる日数をファイルのサイズから割り出し、2種類の作業のそれぞれについて、8名の作業員が一日2時間作業をした場合と4時間作業をした場合で計算をし、前者の場合は、品詞タグで23日程度、文型タグで26日程度かかり、後者の場合は、品詞タグで12日、文型タグで13日かかる計算となった。以下は、データサイズとアルバイトの人員である。

表1 作業のための予測日数と人員

人数	作業内容	総行数	一人当たり	一日当たり	日数
8	POS タグ付け	18822	2352.75	100行(2時間)	23日程度
8	POS タグ付け	18822	2352.75	200行(4時間)	12日程度
8	文構造タグ付け	21151	2643.875	100行(2時間)	26日程度
8	文構造タグ付け	21151	2643.875	200行(4時間)	13日程度

ファイル数 139 ファイル

品詞タグ修正 7名

文型タグ付け 8名

学生アルバイト数 14名（1名品詞修正と文型タグ付与を行なった為重複）

3.1.4 品詞のタグ修正作業内容

品詞のタグデータは以下の点を主に修正した。

- 1) 機械が自動的に付与したタグにエラー（間違った品詞の情報が付与されている）を正しい品詞情報に修正
- 2) 日本語の単語についている品詞タグを修正する：例 `wakaranai_NN1` → `wakaranai_JP`

3.1.5 文型のタグ付け作業

文型のタグ付けは、投野由紀夫教授の作成した文型タグ（資料3）を元にタグを付与してもらった。文型タグでは、単文だけではなく、重文、複文も分析対象とした。統語の発達においては、段階が上がるにつれて複文の産出能力が大きく作用してくるからである。以下、文型タグ付け作業を例示する。

表2 文型タグ付け作業の例:ファイル059

059	And_CC ... the_AT one_PN1 ...	fragments	
059	</s>		
059	<s>		
059	Where_RRQ is_VBZ Sachiko_NP1 tell_VV0 me_PPIO1 !!	SVOO(Wh)	SVR
059	</s>		
059	<s>		
059	Lisa_NP1 said_VVD and_CC Lisa_NP1 's_GE baby_NN1<err crr="went"> go_VV0 </err> poo-poo_JJ and_CC pee-pee_JJ and_CC doo-doo_FU on_II her_PPHO1 ... her_APPGE face_NN1 ...	S and S	SV SVR
059	</s>		
059	<s>		
059	And_CC that_DD1 's_VBZ all_DB ... </J>	SVC(N)	
059	</s>		
059	<s>		
059	<M> ((singing_VVG)) Raindrops_NN2 keep_VV0 fall'in_NN1 on_II my_APPGE head_NN1 ... they_PPHS2 keep_VV0 fall'in_NN1 ...	S and S	SVO(doing) SVO(doing)
059	</s>		
059	<s>		
059	Jun_NPM1 ... </M> <J> ((singing_VVG)) There_EX was_VBDZ a_AT1 farmer_NN1 had_VHD a_AT1 dog_NN1 and_CC Bingo_NN1 was_VBDZ his_APPGE name-oh_UH ...	S and S	There is SVC(N)
059	</s>		
059	<s>		
059	B-I-N-G-O_NN1 ... </J>	fragments	
059	</s>		
059	<s>		
059	<M> Right_RR on_RP ... kid_NN1 </M> <J> Hello_UH ... sir_NN1 ... what_DDQ are_VBR you_PPY do-doing_VDG ...	SVO(N)	
059	</s>		

3.1.6 2012年の活動

2012年は、主に、POSタグの整理と見直しに時間を費やした。POSタグは、基本的なタグ情報の見直しが終わった後、さらに、タグ情報が適切に修正されているかどうかを見直す作業をした。また、話者のタグの付け方に誤りが見つかったためにすべてのファイルのタグを付け直した。

(タグの交差といって、話者情報は必ず<A>内の文字Aが最後にで終る必要があるが、例えば、例えば<A> で始まり で終ったり、後ろが欠けていたりするなどの不備が見つかったために、今度もアルバイト学生の手を借りて、修正を行なった。これには、予想以上に時間がかかったものの、完璧な一致に至ることがなく、何度も見直しが必要となり、そのため本格的な言語分析を開始することが出来ない状態であった。また、全体を見直すためにファイルを一元化して表を作成し、発話の採集についての情報をまとめた。(資料4参照)

3.1.7 2013年度の計画

1) 言語分析

A. 語彙の分析：言語分析ソフトWordSmith version 6.0を使用

語彙の豊富さ (異なり語の豊富さ) Type Token Ratio(TTR)

語彙リストの作成

語彙リストを JACET8000 などの語彙リストと比較する。(その他の語彙リストも検討中である。)

B. 品詞の分析：プログラミング処理を行なう

品詞の種類

各品詞の頻度

C. 句構造の分析：プログラミング処理を行なう（プログラマーに依頼）（予定）

句構造の種類

句構造の頻度

D. 文構造の分析

文型の頻度を計算（単文、複雑文）

E. コミュニケーションストラテジーの分析

コミュニケーションストラテジーについて2系統で分析する（案）

品詞情報を利用して定型表現を抽出して計算する

コミュニケーションストラテジーのリスト（小池作成）を使って

タグを付けて①種類別の頻度、②ストラテジーと文型の相関を時系列で分析

（量的、質的）

2) 分析結果を踏まえて学会口頭発表を行なう。Asia TEFL（2013年10月）に応募する。

3.1.8 2013年の活動：テキストデータ (POSタグの付与なし) を使った語彙の分析

3.1.8.1 学会発表

2013年には、POSタグの整備が完全でなかったために、当面の分析手段としてテキストデータを使った分析を先に行なった。これには、コンピューターソフトWordSmith 言語分析ツールを使用した。WordSmithでは、語彙の異なり語 (TTR: Type Token Ratio) の比較、文の長さの比較 (MLU: Mean Length of Utterances)、語彙リストの作成、特定の単語の入った文の抽出が可能である。そこで、WordSmith を使って、3ヶ月毎の3人の児童の語彙の異なり語と文の平均長を比較した。そこで、分かったことは、語彙の異なりの比率は3人の間でも時間の経過においても、大きな差はなかった。しかし、平均の文の長さを比較すると統計的には3人の間には差がないが、時間が経つと数値が上がっていることが分かった。さらに、語彙の質的な違い語彙リストを作成して比較すると、それぞれの児童は時間が経つにつれて、異なる語彙を増やしているのが分かった。

この結果を踏まえ、マニラで開催されたアジア教師の学会ASIA TEFL INTERNATIONAL CONFERENCEにて発表した。(2013年10月28日) アジアテフルでは、主に以下の3点に絞って発表した。

1) 小池コーパスの背景と意義

2) 語彙の異なり語の児童別、時間軸の比較の結果には、統計的な有意差は無かったが、3ヶ月毎に作成したそれぞれの子どもの語彙リストの語彙の種類が時間の経過と共に豊富になっていることから、3人の語彙の発達が伺える。

3) 3人のこどもの内、年長者の「N」の副詞の発達について。「N」は、3人の中で発話自体は少ないが、automatically, carefully, friendly, usually, mostly, perfectly, partly, typical, information, permission, sabotage, investigatorなどの難易度の高い語彙を産出していることが分かった。

発表内容は小池コーパスの概観を述べると言った程度であったので、詳しい分析は今後の課題となった。

3.1.9 学会発表後の分析

その後の分析 3人の子どもが習得した語彙は、一般の母語話者が平均して使う語彙を程度カバーしているのかを調べた。これには、英語母語話者が一般に使う頻度の高い語彙リストとして知られているGeneral Service List 2000 (West, 1953) と比較をした。3人のこどもの語彙全体とGSLを比較した結果、2284語のうち、782語が同じであることが分かった。これは全体の約34%になる。しかしながら、GSLの語彙リストの1000語以降については、3人のこどもの年齢では日常使わない単語が多く含まれているので、それらを加味すると英語母語話者の一般語彙と小池コーパスの語彙のカバー率はより高くなる可能性がある。例えば、Sachikoの1-3ヶ月の語彙リストとの比較では、発話語彙116個のうち、GSLとの共通語彙は、78個あった。おおまかな計算では、こども達の最初の3ヶ月の語彙の約70%が、英語母語話者が日常使用する語彙を使っていることが分かった。今後はGSLと小池コーパス語彙のより詳しい比較分析を行いたいと考えている。

3.1.10 英語学習者語彙リストとの比較

大学英語教育学会が作成した「JACET8000英単語」(2005)と小池コーパスの語彙リストの比較を行なった。JACET8000語彙はレベルに分かれているので、Level 2までの2000語を小池コーパスと比較することにした。まずは、これらのJACET2000語が英語母語話者の語彙リストであるGSLと比較して、どの程度カバーしているのか(同じ語彙があるのか)についておおまかに把握することにした。比較の結果、2284語のうち、1475語がJACET2000語とGSL語彙リストの共通語彙であることがわかった。つまり、JACET2000語の語彙はGSL語彙の約64.46%をカバーしていることとなる。この数値は高いとは言えないが、半数以上をカバーしている点では、英語母語話者の使用語彙とそれほど離れているとも言えない。今後の分析としては、JACET2000語彙とGSL語彙との共通語を比較しつつ、その中に見られる特徴を探る予定である。

3.1.11 動詞(助動詞、Be動詞を含む)と名詞の発達

テキストデータをもとに、WordSmith を使って、動詞、助動詞、Be 動詞、名詞についてそれぞれの子どもの3ヶ月毎の語彙リストを作成した。その結果動詞については、最初の3ヶ月を越したあたりから豊富な動詞が表れ、過去形の不規則形が規則形よりもやや多めに見られたこと、また、一時的に不規則動詞に“-ED”を付けるエラーがどの子どもにも見られるが、2年目を過ぎるとそういったエラーが激減する。受動態、完了形の文も見られた。助動詞は半年を過ぎたあたりから増える傾向にあり、ほとんどの種類が出そろうが、後半になるとそれほど使われない傾向が見られた。名詞については、どんどん新しい名詞が増える傾向にあるが、これも後半には増える割合が減る傾向にあった。Be動詞については、本動詞と助動詞とあるが、Be 動詞と一般動詞を一緒に使うエラーが後半にまですべてのこどもに見られた。これらは、第二言語習得研究で他の研究にも見られる特徴と似ているものもあるが、日本語の影響も多大であると思われさらなる考察が必要である。全体としては、どんどん語彙が増えていくが最後まで増え続けるのではなく、後半は勢いがなく落ち着いた印象である。これは、別の見方をすれば、英語を話すのに慣れて母語話者のように楽な話し方になっていっていることの表れであるという解釈も出来るだろう。

3.1.12 今後の分析

今後の分析としては、文構造タグのついたデータファイルを集計し、3人のこどもの統語的な発達を調べる計画である。基本的にはPOS品詞タグ付与データ、文型タグデータを中心とした分析を継続して行うことになるが、今回の語彙の分析のようにテキストデータの活用も分析範囲とすることを考えている。

資料-1 発話にタグをつけた資料の一例

<head> (話者、テキストの情報がここから始まる)

<filename>koike013.txt</filename> (ファイルの名前)

<date>1972/12/07</date> (データを採集した日)

<speaker id="S"> (話者の頭文字)

<name>Sachiko</name> (名前)

<sex>female</sex> (性別)

<age>5;03</age> (年齢)

<overseas>0;03;01</overseas> (米国滞在の年月)

<grade></grade>

</speaker>

<speaker id="I">

<name>Paisley</name>

<sex></sex>

<age></age>

<overseas></overseas>

<grade></grade>

</speaker>

<transcriber>Ikuo and Susanne</transcriber> (筆記をした人の名前)

<transdate>1976/02/04</transdate>

<computerized>2005/02/28</computerized> (パソコンに入力した日付け)

<comp_transcriber>Junko Kambara</comp_transcriber> (パソコンに入力した人の名前)

</head> (ここで話者、テキストの情報は終わり)

<body> (ここから本文が始まる)

<S>This is my coat<!--?></!>. Watch out! OK? Come in. Okay. Everybody. Play house! Children. This is, oh!, no, no,no. This man is sister or mommy.</S>

<I>Sister.</I>

<S>Mommy.</S>

<I>Sister.</I>

<S>Mommy. Buiscuit. OK. Fine.</S>

<I>I want to find doll. This doll is sister. And <err crr="they are">this is</err> children. OK.?</I>

<S>How old are you? <!--"/haud ju?/"></!> Knock, knock. Children and brothers.</S>

<I>Five.</I>

<S>Paisley is me. OK.?</S>

<I>OK.</I>

<S>American or Japanese? Oh, no. That's not right. Everybody. Time to play now. Now, book for you, OK.? Monkey and boy. Boy said, "Hi", "Hi". Monkey <err crr="goes">is</err> up, up, up. Monkey <err crr="goes">is</err> round and round and round, 3, 3, 3, and then monkey is up and high. Monkey's

up and high. Everybody comes <err crr=" "></err>and then. Mommy said, "Come here. Get <err crr=" ">the</err> down, down, down <!--minna orite oide></!>. Monkey, everybody good home <ucl></ucl>. Hey.hey,hey. Little girl. Little go. go it. He, he, he, monkey, monkey...<!--error?>said</!>. Momma going, jump, jump, jump.</S>
<I>...go to....</I>

資料2 CLAWS7 品詞のタグセット(右のコラムは左のコラムの説明)

UCREL CLAWS7 Tagset	
APPGE	possessive pronoun, pre-nominal (e.g. my, your, our)
AT	article (e.g. the, no)
AT1	singular article (e.g. a, an, every)
BCL	before-clause marker (e.g. in order (that),in order (to))
CC	coordinating conjunction (e.g. and, or)
CCB	adversative coordinating conjunction (but)
CS	subordinating conjunction (e.g. if, because, unless, so, for)
CSA	as (as conjunction)
CSN	than (as conjunction)
CST	that (as conjunction)
CSW	whether (as conjunction)
DA	after-determiner or post-determiner capable of pronominal function (e.g. such, former, same)
DA1	singular after-determiner (e.g. little, much)
DA2	plural after-determiner (e.g. few, several, many)
DAR	comparative after-determiner (e.g. more, less, fewer)
DAT	superlative after-determiner (e.g. most, least, fewest)
DB	before determiner or pre-determiner capable of pronominal function (all, half)
DB2	plural before-determiner (both)
DD	determiner (capable of pronominal function) (e.g any, some)
DD1	singular determiner (e.g. this, that, another)
DD2	plural determiner (these,those)
DDQ	wh-determiner (which, what)
DDQGE	wh-determiner, genitive (whose)
DDQV	wh-ever determiner, (whichever, whatever)
EX	existential there
FO	formula
FU	unclassified word
FW	foreign word
GE	germanic genitive marker - (' or's)
IF	for (as preposition)

資料3 文型タグをつけるためのタグセット

単一文型	例
SV	Birds fly.
SVP	Our car is in the garage.
SVR	I'm here
SVC(N)	This is my computer. / Mary became a nurse.
SVC(J)	She is pretty.
SVC(wh)	The question is whether he will come or not.
SVC(that)	The point is that we're not going there.
SVC(doing)	My hobby is watching English drama on TV.
it ... that ...	It is obvious that John is honest.
SVO(N)	I play baseball. / Jim has a good car.
SVO(wh)	I don't know whether he will come or not.
SVO(to do)	I want to have something cold to drink.
SVO(doing)	Boys like playing baseball.
SVO(that)	I think that John is a liar.
SVOP	John put the key in the lock.
SVOO	John gave me a book.
SVOO(wh)	He taught me how to play the guitar.
SVOO(that)	She told us (that) the story was true.
SVOO(to do)	I told John to see a doctor.
SVOC(N)	Everyone calls me Bob.
SVOC(J)	I found the book very interesting.
SVOC(doing)	He saw the wood burning.
SVOC(do)	I made her go against her wil.
SVOC(done)	I had my purse stolen.
SVOC(to do)	Do you want me to help? / I thought John to be honest.
passive	English is spoken here.
重文	
S and S	S 内は上記の単一文型で整理
S but S	
S or S	
複文	
When S, S	他の副詞節も同様 (so, because, as, etc.)
If S, S	
S-RC	関係節の埋め込み文
that 節	上記の分類を与えた後、that 内の文型も記述
wh 節	上記の分類を与えた後、that 内の文型も記述

3.2 日本人児童の英語発話データに見られる語彙の発達

はじめに

この研究は、日本人児童3名の英語発話コーパスを使って、彼らの語彙の習得の軌跡を調査分析したものである。英語発話コーパスは、Koike(1983)の3人の日本人の子どもの英語による発話をパソコンのテキスト文書として保存したもので、「小池コーパス」と名称を付けた。この発話データは、Koike(1983)の博士論文のために採集されたもので、1972年から1975年の間に約2週間に一度定期的に採集されたもので、1) 第二言語としての英語の長期発話データであること 2) 対象者が日本人のこどもであることのいずれにおいても研究的価値の高い存在である。第二言語としての英語のデータは多く存在するが、長期的なデータとなると希少であり、さらに日本人による長期的な英語の発話データとなると、筆者の調べる限りでは Hakuta (1974) による Uguisu という5才の日本人の子どもの発話を長期に渡って記録したものがのみである。一方で、最近ではコーパス言語学の興隆に伴い日本においても「日本人1200人の英語スピーキングコーパス」(和泉絵美他編、2004)「日本人中学生1万人の英語コーパス」(投野由紀夫編、2007)などに代表されるように日本人の英語学習者の発話やライティングの大規模なデータを使った言語分析・言語習得研究が盛んになってきている。小池コーパスの言語分析においては、このようなコーパス言語学の発達の恩恵を受け、過去の分析に加えコーパス的な分析手法を含んだ言語分析として第二言語習得研究分野とコーパスを使った言語研究分野への貢献度は大きいと思われる。

3.2.1 データについて

3.2.1.1 協力者

データ協力者は日本人の児童3名であり、小池生夫氏の子弟である。彼らは、1972年秋に渡米し両親の元で地元の小学校の通常クラスへ通学した。従って、言語習得的に言えば、第二言語を自然に習得する環境である。以下は3人の子どものデータ採取開始時の年齢である。「;」は「年」を意味する。開始は、彼らが米国に住み始めて2ヶ月後である。データの整理上、見やすいように頭文字で表記をした。以下は「S」「J」「N」とする。

「S」	(女兒)	5;4
「J」	(男児)	7;8
「N」	(男児)	11;2

3.2.1.2 データの形式

小池コーパスの形態を細かく述べると、現時点ではデータは次の3つの形式で保存されている。1) テキスト形式(パソコン内のメモ帳) 2) 品詞タグを付与したデータ(パソコン内のエクセルシート)、3) 文構造タグを付与したデータ(パソコン内のエクセルシート)、である。この研究においては、1)のテキスト形式のデータを使用した。それは、今回主に使用する語彙分析ツール WordSmith というソフトに対してテキスト形式が適しているからである。データの数、種類、量は以下の通りである。以下のリストで示したように、全ファイルは139あるが、分析に適さない歌、暗唱、英文和訳などは除外したので、今回の分析では、131ファイルを使用した。表2で示しているように、発話を抽出するためのタスクは複数あり、主なものとして、英語話者が子ども達に話しかけるもの(インタビュー)、会話(子ども同士、または子どもと友人)、絵本の再話(「シ

ンデレラ」「赤頭巾」「ピノキオ」など)、絵本などの描写などがある。表1はタスクの数と主な6タスクの説明である。

- ・データ採集時期 1972年10月～1975年4月(約2年半)
- ・データ採集回数 定期的に週2回の頻度
- ・データファイル数 139(一回の採集を1ファイルと数える)
- ・分析に使用したファイル数 131ファイル

表1 タスクの種類と数

タスクの種類	数	説明
会話	51	子ども同士の会話
Interview	37	英語話者と子ども達との会話
ストーリー再話	22	童話の再話
描写	10	絵や写真の描写
和文英訳	6	日本語を英語に訳す
モノローグ	4	子ども一人の発話
Reading & Transcription	2	子どもが絵本を見て英語にする
合計	132	

* 歌、暗唱、英文和訳は含めない

* ひとつのファイルに2種類のものを一つ含む

3.2.1.3 語彙の分析方法

語彙の分析については、以下のリサーチクエッションを立てた。

- (1) こども3人の語彙の発達に差はあるのだろうか。
- (2) こども3人の名詞、動詞(助動詞、Be動詞を含む)はどのように発達しているのか。
- (3) (2)について、エラーにはどのような特徴があるだろうか。

これらの分析には、言語分析ツールであるWordSmith (version 6.0)を使用した。これは、多くの機能を持つ言語分析ツールであるが、今回の分析に関連して言えば、テキストファイルに使われる語彙を抽出して、語彙リストを作成する機能を持っている。この語彙リストでは、異なり語の数、とデータファイルの総単語数を調べてくれる。また、キーワードの検索によって、語彙リスト内の特定の語彙の元のテキストを見ることが出来るので、単語がどのような使われ方をしているのかを知るために便利である。

3人の語彙の発達について、以下のように2つのリサーチクエッションを立てた。

- (1) RS1: 3ヶ月ごとのデータ間における3人の語彙の種類に違いはあるか?

WordSmithを使い、3人のテキストデータを3ヶ月毎に分け、そのテキストの中でどのような種類の単語が使われているのかを調べた。

- (2) RQ2: 語彙の中で名詞、動詞、機能語はどのように発達しているのか?

(1)で作成した語彙リストの中から名詞、動詞、助動詞のみ抽出し、特に進行形、過去形(規則形、不規則形、3単元の-sがどれくらい使われるのか調べた。これについては、「S」の語彙の発達を中心に調べた。

「S」は3人の中で最も年少である。

3.3 結果

3.3.1 3人のこどもの発話語彙の比較

はじめに、全体の語彙についての様相をつかむために3ヶ月毎の語彙数と、コーパス全体の語彙数をWordSmithを使って計算した。下の表2は、3人のこどもの異なり語と発話の総語数を3ヶ月毎に計算したものである。表2の縦軸の数字はSTTRの比率を表し、横軸の数字は3ヶ月おきの経過月数である。多少の違いはあるが、3人とも時間の経過と共に、語彙数が増えているのがわかる。同時に、全体から見ると、22ヶ月あたりから語彙量の増加はなくなってきている。つまり、語彙をどんどん覚える時期があるが、それがずっと続くということはないことを示している。

表2 S, J, N の発話語彙数 (3ヶ月単位)

		1-3	4-6	7-9	10-12	13-15	16-18	19-21	22-24	25-28	29-31	32-34
S	異なり語	121	362	408	843	915	699	701	673	561	357	388
	総語彙数	286	1438	1633	5706	6681	3575	2976	2811	2256	1310	1277
J	異なり語	174	183	507	623	842	714	524	334	587	239	345
	総語彙数	303	442	1563	3201	4846	3689	2159	936	2526	772	1148
N	異なり語	191	182	526	691	567	483	461	233	290	44	145
	総語彙数	387	424	1882	3443	2359	2183	1662	551	708	55	392

注) 異なり語：発話の中の異なる語彙。GO が何回表れても1回と数えるが、GO, GOES で形が違う語彙と考え2つとカウントする。

3.3.2 助動詞、動詞、名詞の発達

次に、動詞、助動詞、名詞の発達を見る。動詞については、異なり語をすべてカウントしたもの（例えば、LIVE, LIVING, LIVESで3つとする）と規則変化をする動詞の場合は見出し語は一つであるが、不規則変化をするものは、別の語彙としてカウントした（例えば、GO, WENT, GONEは別の語彙とカウント）。助動詞は、“won’ t” “wouldn’ t” などの縮約の否定文も習得語彙に入れて分析した。また、例えば、I dunno. (=I don’ t know) の“dunno” を一つの単語として使用しているように思われるからである。これは自然な言語習得において、当然起こることであるが、耳から固まりで入った音と意味をつなげて言語を習得する過程において、縮約しているという意識もなく一つの「単語」として使うようである。全体としては、表1から伺えるように、3人のこどもの最初の3ヶ月の助動詞、動詞、は少ないが、次の4-6ヶ月には「N」以外は急激に増加する。「N」の語彙はその後他の二人と並んで増えていくが、発話がスムーズに出るまでには、他の二人よりは時間がかかっている。これには、11歳という年齢の精神状態も関与しているかも知れない。

以下の表3は「S」による助動詞の発達の推移である。最初の3ヶ月は、“Do”, “Don’ t” のみであるが、徐々に種類が増え、10-12ヶ月では過去形、否定形縮約形(“Won’ t” など)が見られる。ここで、全体としての特徴として、10ヶ月から12ヶ月と、16ヶ月から18ヶ月が最も助動詞の種類が多い。助動詞のhad, have, has は、殆どが“have to” で使われていた。3つの中でhad, has 次第に使われるようになるが、次第に数が減り、最後に“have to” は残って使われているようであった。

表3 Sの助動詞の発達の推移

1-3	4-6	7-9	10-12	13-15	16-18	19-21	22-24	25-28	29-31	32-34
DO	CAN	CAN	CAN	CAN	CAN	CAN	CAN	CAN	CAN	CAN
DONT	CAN'T	CAN'T	CAN'T	CAN'T	CANNOT	CAN'T	CANNOT	CAN'T	CAN'T	DID
	DID	COULD	COULD	COULD	CAN'T	COULD	COULD	COULD	COULD	DO
	DO	DID	COULDN'T	DID	COULD	DIDN'T	DID	DID	COULDN'T	DON'T
	WONT	DIDN'T	DID	DIDN'T	DID	DOES	DIDN'T	DIDN'T	DID	HAVE
	(WILL)'LL	DO	DIDN'T	DOES	DIDN'T	DOESN'T	DOES	DO	DIDN'T	MAY
		DON'T	DOES	DOESN'T	DO	DON'T	DOESN'T	DOES	DO	MIGHT
		HAS	DOESN'T	DON'T	DOESN'T	MAY	DON'T	DOESN'T	DOES	MUST
		HAVE	DON'T	HAS	DON'T	MIGHT	HAD	DON'T	DOESN'T	SHALL
		MAY	HAD	HAVE	HAS	MUST	HAS	DONTYOU	DON'T	WILL
		MUST	HAS	MIGHT	HAVE	SHOULD	HAVE	HAD	HAS	WOULD
		SHOULD	HAVE	WILL	HAVE'T	WOULD	MAY	HAS	MIGHT	(WILL)'LL
		(WILL)'LL	HAVEN'T	WON'T	MAY	WOULDN'T	MIGHT	HAVE	SHOULD	(WOULD)'D
			MAY	WOULD	MIGHT	(WILL)'LL	SHOULD	MIGHT	WILL	(HAVE)'VE
			MIGHT	(WILL)'LL	MUST	(HAVE)'VE	WILL	MUST	WON'T	DON'TCHA
			SHALL		SHALL		WON'T	SHOULD	WOULD	
			SHOULD		WILL			WILL	(HAVE)'VE	
			WEVE		WON'T			WON'T		
			WILL		WOULD			WOULD		
			WON'T		YOU'VE			(WILL)'LL		
			WOULD		(WILL)'LL			(WOULD)'D		
			(WILL)'LL		(WOULD)'D					
			(WOULD)'D		DYOU					

次に動詞の発話数と、過去形の不規則変化、規則変化を見る。以下の表4はSの動詞の数と発話された変化形の数をまとめたものである。動詞の数は、10ヶ月を過ぎるとかなり多くだされているのが分かる。よく見ると、最初に進行形のINGが使われだすが、過去の動詞の不規則形が先に現れるのが興味深い。規則形の過去は、3単元の-Sよりも少し遅れてよく使われるようになるようであった。また、13ヶ月以降の一時期不規則動詞にEDをつけるエラーが見られ、時期が過ぎると殆どなくなる傾向になったことは注目される。これは、英語の動詞の文法的規則がしばらく英語に触れているうちに内在化して、無意識にルールを(間違っ)使うという普遍的なエラーであり、自然に言語を習得する過程において重大なプロセスである。

もうひとつの特徴は、以下の例1のような、BE動詞と一般動詞を同時に使うエラーである。このタイプのエラーは、中間言語としてしつこく残る傾向にあるようである。これは、Dady is coming. の不完全な文という解釈も出来るし、Daddy came in. と言おうとして、BE動詞をつけてしまったということも考えられる。

例1: *Daddy is come in. (文脈からDaddy came in.が正しい文)

表4 Sの動詞の変化形の発達の推移

	1-3	4-6	7-9	10-12	13-15	16-18	19-21	22-24	25-28	29-31	32-34
異なり語	8	77	72	175	186	180	135	101	125	96	78
見出し語	8	66	61	137	129	134	106	82	102	78	68
ING	0	9	12	31	44	33	15	15	10	8	7
規則過去(ED)	0	0	2	9	20	23	17	10	17	4	11
不規則過去	1	4	6	12	16	14	16	10	20	11	0
規則過去分詞	0	1	0	1	2	0	0	0	1	0	0
不規則過去分詞	0	1	0	1	2	2	2	0	1	1	12
3単現のS	0	4	6	10	17	15	18	13	8	9	4

次の表5と表6はそれぞれ、Sの発話に見られる動詞のリストで、表5は1-9ヶ月、表6は22-28ヶ月である。語彙の質は基本的に似通っているが、表6のリスには、不規則過去形、3単元の-S、規則変化の過去形が多く見られる。

表5 Sの発話にある動詞リスト(1-9ヶ月)

1-3			4-6			7-9		
Word	頻度	%	Word	頻度	%	Word	Freq.	%
DO	6	2.00	LIKE	21	1.46	LIKE	29	0.18
WELCOME	2	0.67	COME	19	1.32	HAVE	19	0.31
THANK	2	0.67	PLAY	12	0.83	KNOW	17	0.06
SIT	2	0.67	SAID	11	0.76	SEE	14	0.06
GO	2	0.67	GO	9	0.62	GO	11	0.12
PLAY	1	0.33	HAVE	9	0.62	PUT	10	0.06
COME	1	0.33	KNOCK	8	0.56	SAID	8	0.06
CAME	1	0.33	KNOW	8	0.56	PRETEND	8	0.06
			PUT	8	0.56	GET	7	1.78
			SEE	8	0.56	COME	7	1.78
			WANT	8	0.56	WEAR	6	0.06
			GET	6	0.42	EAT	6	0.24
			TAKE	6	0.42	BROKE	6	0.24
			GONNA	5	0.35	USE	5	0.06
			DRAW	5	0.35	SAY	5	0.06
			DO	5	0.35	PLAY	5	0.06
			HURRY	5	0.35	GONNA	5	0.31
			LOOK	5	0.35	BRING	5	0.31
			SAY	5	0.35	SLEEP	4	0.12
			STUDY	4	0.28	GOT	4	0.49
			SLEEPING	3	0.21	BRUSH	4	0.18
			LET'S	3	0.21	THINK	3	0.06
			FORGET	3	0.21	TELL	3	0.37
			GIVE	3	0.21	STEP	3	0.06
			KICK	3	0.21	MAKE	3	0.06
			MAKE	3	0.21	GIVE	3	0.18
			SIT	3	0.21	COMING	3	0.37
			STAND	3	0.21	CLEAN	3	0.43
			THANK	3	0.21	TALK	2	0.06
			WAIT	3	0.21	MOVE	2	0.18
			COMING	2	0.14	GOING	2	0.06
			GOING	2	0.14	FORGOT	2	0.06
			PLAYING	2	0.14	END	2	0.06
			STUDYING	2	0.14	DOING	2	0.18
			GAVE	2	0.14	CRYING	2	0.06
			EAT	2	0.14	CRY	2	0.12
			NEED	2	0.14	WORK	1	0.06
			SMELL	2	0.14	WANT	1	0.06
			STEP	2	0.14	WALKING	1	0.06
			TALK	2	0.14	WALK	1	0.06
			THINK	2	0.14	THREW	1	0.49
			TIE	2	0.14	THANK	1	0.12
			USE	2	0.14	TAKE	1	0.06
			DOING	1	0.07	STUDY	1	0.06
			MAKING	1	0.07	STAPLED	1	0.12
			SHOPPING	1	0.07	SPEND	1	0.06
			TURNING	1	0.07	SLEEPS	1	0.06
			GOT	1	0.07	SLEEPING	1	0.12
			FINISHED	1	0.07	SIT	1	0.18
			STUCK	1	0.07	SAYS	1	0.06
			COMES	1	0.07	SAW	1	0.31
			GOES	1	0.07	PUSHED	1	0.06
			HAS	1	0.07	PLAYING	1	0.37
			LIKES	1	0.07	NOTICE	1	0.06
			COOK	1	0.07	NEED	1	0.31
			DRINK	1	0.07	MADE	1	0.18
			HEAR	1	0.07	LOOKS	1	0.12
			HIT	1	0.07	LIVING	1	0.06
			LEAVE	1	0.07	LIVES	1	0.06
			LET	1	0.07	LET	1	0.06
			PAY	1	0.07	LEARN	1	0.24
			PRETEND	1	0.07	JUMP	1	0.86
			READ	1	0.07	GOTTA	1	0.61
			RETURN	1	0.07	GOES	1	0.06
			RIDE	1	0.07	FORGET	1	0.06
			RUN	1	0.07	FIXING	1	1.04
			SCRATCH	1	0.07	EATING	1	1.16
			SELL	1	0.07	DONNO	1	0.67
			SHOW	1	0.07	COLORING	1	0.12
			SING	1	0.07	BEING	1	0.06
			SLEEP	1	0.07	DO	1	0.43
			SPANK	1	0.07			
			SPEAK	1	0.07			
			TURN	1	0.07			
			WALK	1	0.07			
			WATCH	1	0.07			
			WEAR	1	0.07			

表6 Sの発話に見られる動詞リスト

22-24			25-28		
Word	Freq.	%	Word	Freq.	%
KNOW	16	0.68	SEE	16	1.22
GO	15	0.64	KNOW	14	1.07
LIKE	14	0.64	LIKE	14	1.07
PUT	14	0.60	COME	12	0.92
GET	12	0.51	GET	11	0.84
DO	10	0.43	PUT	11	0.84
SEE	10	0.43	WANNA	8	0.61
WENT	9	0.38	GO	8	0.61
REMEMBER	8	0.34	GONNA	6	0.46
KNOCK	7	0.30	LOOK	6	0.46
PRETEND	7	0.30	PRETEND	5	0.38
FORGOT	6	0.26	NEED	4	0.31
LET'S	6	0.26	SWIM	4	0.31
HAS	6	0.26	TAKE	4	0.31
COME	6	0.26	THOUGHT	3	0.23
THINK	6	0.26	LOOKS	3	0.23
FIT	6	0.21	LIVE	3	0.23
LOOK	5	0.21	MOVE	3	0.23
PLAY	5	0.21	OPEN	3	0.23
SIT	5	0.21	RIDE	3	0.23
WANT	5	0.21	TELL	3	0.23
TRIED	4	0.17	THINK	3	0.23
TURNED	4	0.17	WANT	3	0.23
WATCHING	4	0.17	HAD	2	0.15
CAME	4	0.17	GOING	2	0.15
SAID	4	0.17	MISSING	2	0.15
TAKE	4	0.17	FIXED	2	0.15
WAIT	4	0.17	HAPPENED	2	0.15
INVITED	3	0.13	HOLD	2	0.15
WANTED	3	0.13	HURT	2	0.15
FELL	3	0.13	KEEP	2	0.15
GOT	3	0.13	BRING	2	0.15
HAD	3	0.13	MEET	2	0.15
LEFT	3	0.13	SIT	2	0.15
SITS	3	0.13	SLEEP	2	0.15
EAT	3	0.13	STAND	2	0.15
FIX	3	0.13	TURN	2	0.15
LEAVE	3	0.13	USE	2	0.15
SAY	3	0.13	WORK	2	0.15
WORK	3	0.13	COMES	1	0.08
DANCED	2	0.09	DRINKS	1	0.08
FIXING	2	0.09	FITS	1	0.08
MADE	2	0.09	GOES	1	0.08
SAT	2	0.09	HAS	1	0.08
SAW	2	0.09	SAYS	1	0.08
TOOK	2	0.09	SLEEPS	1	0.08
DONE	2	0.09	THINKS	1	0.08
MARRIED	2	0.09	GONE	1	0.08
SHINES	2	0.09	LET'S	1	0.08
WANNA	2	0.09	BOUGHT	1	0.08
ASK	2	0.09	BROKE	1	0.08
DRINK	2	0.09	BROUGHT	1	0.08
FIND	2	0.09	CAME	1	0.08
HELP	2	0.09	FORGOT	1	0.08
HOPE	2	0.09	GOT	1	0.08
MEET	2	0.09	SAID	1	0.08
SPEAK	2	0.09	TOLD	1	0.08
STAY	2	0.09	WENT	1	0.08
TRY	2	0.09	COMING	1	0.08
WAKE	2	0.09	DOING	1	0.08
WEAR	2	0.09	FALLING	1	0.08
ALLOWED	1	0.04	FROSTING	1	0.08
APPEARED	1	0.04	LIVING	1	0.08
CALLED	1	0.04	SWIMMING	1	0.08
CLIMBED	1	0.04	TAKING	1	0.08
CRIED	1	0.04	BUMFED	1	0.08
FITTED	1	0.04	TICKLED	1	0.08
HAPPENED	1	0.04	FIND	1	0.08
LIVED	1	0.04	FIX	1	0.08
MOVED	1	0.04	FORGET	1	0.08
PICKED	1	0.04	HIT	1	0.08
PLAYED	1	0.04	HURRY	1	0.08
STAYED	1	0.04	ASK	1	0.08
STOPPED	1	0.04	BREAK	1	0.08
STRIKED	1	0.04	BREATHE	1	0.08
WATCHES	1	0.04	BUMP	1	0.08
GOTTA	1	0.04	CLEAR	1	0.08
COMING	1	0.04	CLOSE	1	0.08
DANCING	1	0.04	DIVE	1	0.08
HAVING	1	0.04	DRINK	1	0.08
LOOKING	1	0.04	FALL	1	0.08
MOVING	1	0.04	LISTEN	1	0.08
PLAYING	1	0.04	PLAY	1	0.08
SHOPPING	1	0.04	PRESS	1	0.08
WASHING	1	0.04	RUN	1	0.08
ATE	1	0.04	SHUT	1	0.08
BOUGHT	1	0.04	TALK	1	0.08
BROKEN	1	0.04	TRY	1	0.08
GAVE	1	0.04	WAIT	1	0.08
RODE	1	0.04	WISH	1	0.08
STRUCK	1	0.04	WONDER	1	0.08
THOUGHT	1	0.04	WORRY	1	0.08
WOKE	1	0.04	HAVING	1	0.08
LOOKS	1	0.04	DO	1	0.08
PUTS	1	0.04	HAVE	1	0.08
TALKS	1	0.04			
WORKS	1	0.04			
BEND	1	0.04			
BRING	1	0.04			
CALL	1	0.04			
COOK	1	0.04			
CRASH	1	0.04			
DIE	1	0.04			
FALL	1	0.04			
FASTEN	1	0.04			
FOUND	1	0.04			
GIVE	1	0.04			
HIDE	1	0.04			
LET	1	0.04			
LISTEN	1	0.04			
MIND	1	0.04			
NEED	1	0.04			
PACK	1	0.04			
RAN	1	0.04			
SHARE	1	0.04			
START	1	0.04			
SURPRISE	1	0.04			
TALK	1	0.04			
USE	1	0.04			
WANTS	1	0.04			
WASH	1	0.04			
WATCH	1	0.04			

次にSの名詞の発達を動詞と比べて見る。以下の表7は、Sの名詞と動詞の3ヶ月毎のデータに見られる異なり語(種類)である。全体として名詞の数が多い。これは、言語が動詞を中心として、名詞やその他の品詞から成り立っていることを考えれば当然であるが、名詞の数が動詞の数が増えるのと同時に大量の名詞が使われだしている。しかし表6でも分かるように、全体の総異なり語数との割合を見ると、平均して、名詞は20%~30%弱、動詞はそれよりやや低くしかし、名詞の比率と大きな差はなく、一定しているのが分かる。図1はそれを折れ線グラフで表したものである。最初の名詞の比率が高いのは、タスクで物の名前を聞いて答えるため名詞の数が多くなっているためである。しかし、表7を見ると、全体として動詞がよく使われていることがわかるのではないだろうか。

表7 Sの名詞と動詞の発達の比較

	1-3	4-6	7-9	10-12	13-15	16-18	19-21	22-24	25-28	29-31	32-34
名詞の異なり語	80	108	125	242	277	183	184	129	152	75	106
動詞の異なり語	8	77	72	175	186	180	135	101	125	96	78
ファイル内異なり語	121	362	408	843	915	699	701	673	561	357	388
総発話語彙数	286	1438	1633	5706	6681	3575	2976	2811	2256	1310	1277

表8 Sの全体の異なり語に対する名詞と動詞の異なり語の割合

	1-3	4-6	7-9	10-12	13-15	16-18	19-21	22-24	25-28	29-31	32-34
名詞の異なり語	66%	41%	31%	29%	26%	26%	26%	19%	27%	21%	27%
動詞の異なり語	7%	21%	18%	19%	20%	26%	19%	15%	22%	27%	20%

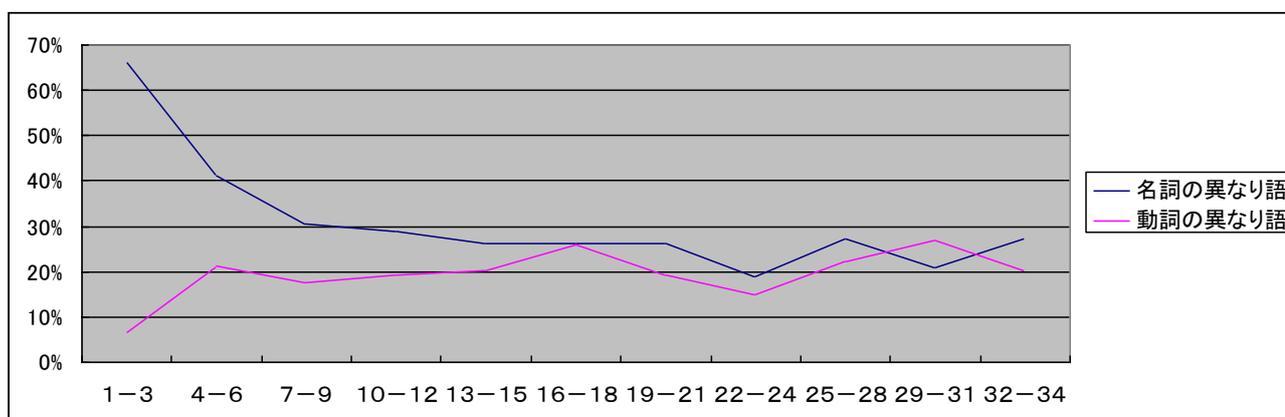


図1 Sの全体の異なり語と名詞と動詞の異なり語との比較

表9はSの発話名詞リストである。単語は頻度順に並んでいる。単語の横のコラムは頻度数を表し、次のコラムは全テキストに対する頻度のパーセンテージである。名詞はタスクによって左右されるが、最初は日常生活についての質問にこたえるタスクが多いので、学校での話や、友達と何をして遊ぶかなどについてのトピックに合わせた単語が使われている。

表9 Sの発話にある名詞リスト（1ヶ月から9ヶ月まで）

1-3			4-6			7-9		
単語	頻度	%	単語	頻度	%	単語	頻度	%
DOG	8	2.67	MONKEY	19	1.32	ZOO	2	0.12
FIVE	6	2.00	ONE	14	0.97	YEAR	1	0.06
ORANGE	4	1.33	BABY	6	0.42	WOLF	4	0.24
PEA	4	1.33	TEACHER	6	0.42	WILLOW	1	0.06
BANANA	3	1.00	BOY	5	0.35	WHITE	4	0.24
BIRD	3	1.00	CANDLE	5	0.35	WAY	2	0.12
DADDY	3	1.00	GIRL	5	0.35	WATER	2	0.12
DANCE	3	1.00	HOUSE	5	0.35	WALL	2	0.12
DENTIST	3	1.00	TODAY	5	0.35	UNCLE	1	0.06
DOCTOR	3	1.00	BIRTHDAY	4	0.28	TWO	1	0.06
HELICOPTER	3	1.00	BOOK	4	0.28	TWELVE	1	0.06
ICECREAM	3	1.00	ENGLISH	4	0.28	TIME	6	0.37
MILK	3	1.00	KIND	4	0.28	THREE	1	0.06
NAME	3	1.00	RIGHT	4	0.28	THIRTY	1	0.06
THREE	3	1.00	BED	3	0.21	THING	2	0.12
ZEBRA	3	1.00	FLOWER	3	0.21	THEATER	1	0.06
APPLE	2	0.67	JAPANESE	3	0.21	TEETH	4	0.24
BACK	2	0.67	KIDS	3	0.21	TEACHER	1	0.06
BALLOON	2	0.67	PANDA	3	0.21	SUNDAY	1	0.06
BLACK	2	0.67	PIECE	3	0.21	SUMMER	1	0.06
BOWL	2	0.67	RED	3	0.21	SPRING	2	0.12
BOY	2	0.67	SISTER	3	0.21	SONG	2	0.12
BROTHER	2	0.67	THINGS	3	0.21	SHOES	1	0.06
BUTTER	2	0.67	BLACK	2	0.14	SECRET	2	0.12
CRAYON	2	0.67	CHILDREN	2	0.14	SCHOOL	4	0.24
DESK	2	0.67	DADDY	2	0.14	ROOM	1	0.06
DOOR	2	0.67	DOLL	2	0.14	RED	1	0.06
EIGHT	2	0.67	DRESS	2	0.14	RABBIT	1	0.06
ELEPHANT	2	0.67	INDIAN	2	0.14	QUEEN	2	0.12
EYE	2	0.67	MOM	2	0.14	PLATE	1	0.06
FLOWER	2	0.67	NAME	2	0.14	PIZZA	2	0.12
GIRL	2	0.67	PART	2	0.14	PINK	2	0.12
GREEN	2	0.67	PARTY	2	0.14	PILLLOW	3	0.18
HAIR	2	0.67	RECORDER	2	0.14	PIECE	1	0.06
HAT	2	0.67	SCHOOL	2	0.14	PICTURE	1	0.06
HOUSE	2	0.67	SHOES	2	0.14	PEOPLE	6	0.37
IRON	2	0.67	SONG	2	0.14	PENCIL	1	0.06
KITE	2	0.67	TAPE	2	0.14	PEE	2	0.12
KNIFE	2	0.67	TIME	2	0.14	PARK	1	0.06
LOOP	2	0.67	TREE	2	0.14	PANTS	6	0.37
MAMMY	2	0.67	WAY	2	0.14	PAJAMAS	1	0.06
MOON	2	0.67	WOLF	2	0.14	ONE	10	0.61
MORNING	2	0.67	HUSBAND	1	0.07	NOSE	1	0.06
MOUTH	2	0.67	ALINT	1	0.07	NINE	1	0.06
NIGHT	2	0.67	BACK	1	0.07	NIGHT	2	0.12
NINE	2	0.67	BEADS	1	0.07	NAME	1	0.06
ONE	2	0.67	BLUE	1	0.07	MUSHROOM	2	0.12
OPEN	2	0.67	BODY	1	0.07	MOVIE	2	0.12
PANTS	2	0.67	BUISCUIT	1	0.07	MOTHER	3	0.18
PEANUTBUTT	2	0.67	CAKE	1	0.07	MORNING	4	0.24
PIANO	2	0.67	CANDY	1	0.07	MONKEY	1	0.06
PICTURE	2	0.67	CAR	1	0.07	MOMMIES	1	0.06
PINK	2	0.67	CAROL	1	0.07	MIDDLE	2	0.12
PURPLE	2	0.67	CHURCH	1	0.07	MEAT	1	0.06
RED	2	0.67	CLOCK	1	0.07	MAMA'S	1	0.06
SEVEN	2	0.67	CLOTHES	1	0.07	LUNCH	3	0.18
SIX	2	0.67	COAT	1	0.07	LIVINGROOM	1	0.06
SKI	2	0.67	COOKIES	1	0.07	LIGHT	1	0.06
SOCKS	2	0.67	DAY	1	0.07	LADY	1	0.06
SOUP	2	0.67	DOGGIE	1	0.07	KITCHEN	9	0.55
TEN	2	0.67	DON	1	0.07	KINDERGART	2	0.12
TOMATO	2	0.67	DUCK	1	0.07	KIND	2	0.12
TURTLE	2	0.67	ERASER	1	0.07	JUICE	1	0.06
TWO	2	0.67	FACE	1	0.07	JAPANESE	2	0.12
WATER	2	0.67	FAN	1	0.07	JAPAN	1	0.06
WHITE	2	0.67	FIFTEEN	1	0.07	INDIAN	1	0.06
YELLOW	2	0.67	FISH	1	0.07	INCHES	1	0.06
AFTERNOON	1	0.33	FORTEEN	1	0.07	HOUSE	7	0.43
ALLIGATOR	1	0.33	FRIEND	1	0.07	HOOD	1	0.06
BARBAR	1	0.33	GUEST	1	0.07	HOME	2	0.12
CANARY	1	0.33	HAT	1	0.07	HAND	1	0.06
CHICKEN	1	0.33	JACK'OLANTE	1	0.07	HAM	1	0.06
COMB	1	0.33	JAPAN	1	0.07	GUEST	1	0.06
EAR	1	0.33	KEY	1	0.07	GREEN	1	0.06
FOUR	1	0.33	KITCHEN	1	0.07	GRASS	1	0.06
MAN	1	0.33	LADY	1	0.07	GRANDMOTH	4	0.24
MONKEY	1	0.33	MAMA	1	0.07	GRANDFATHE	1	0.06
SANDWITCH	1	0.33	MAMMY	1	0.07	GIRL	1	0.06
SLAID	1	0.33	MAN	1	0.07	GIPPER	1	0.06
YEARS	1	0.33	MINUTE	1	0.07	GHOST	1	0.06
			MUSIC	1	0.07	FRONT	1	0.06
			ORANGE	1	0.07	FRIEND	1	0.06
			PAIR	1	0.07	FORTY	1	0.06
			PAPA	1	0.07	FLOWERS	2	0.12
			PAPERS	1	0.07	FLOOR	3	0.18
			PURSE	1	0.07	FISH	2	0.12
			POURANGE	1	0.07	FEET	1	0.06
			RECORD	1	0.07	EYES	1	0.06
			ROBBERS	1	0.07	ENGLISH	1	0.06
			SEA	1	0.07	ELEPHANTS	1	0.06
			SEVEN	1	0.07	EGG	2	0.12
			SLEIGH	1	0.07	EASTER	1	0.06
			SWEATER	1	0.07	DUTCH	1	0.06
			TEN	1	0.07	DRESS	3	0.18
			TWELVE	1	0.07	DOLLS	3	0.18
			TWILIGHT	1	0.07	DOG	1	0.06
			WATER	1	0.07	DECORATION	1	0.06
			WINDOW	1	0.07	DAY	4	0.24
			WORD	1	0.07	DANCE	1	0.06
			YELLOW	1	0.07	DAD	1	0.06
					CUP	1	0.06	
					COCK	2	0.12	
					CLOTHES	2	0.12	
					CLOCK	1	0.06	
					CHURCH	1	0.06	
					CHOCOLATE	2	0.12	
					CHERRY	1	0.06	
					CENTIMETER	2	0.12	
					CHEESECAK	1	0.06	
					CAKE	2	0.12	
					BUS	1	0.06	
					BOY	1	0.06	
					BLUE	1	0.06	
					BLOSSOMS	1	0.06	
					BIRTHDAY	1	0.06	
					BED	2	0.12	
					BEARS	1	0.06	
					BALL	3	0.18	
					BAG	1	0.06	
					BABY	6	0.37	
					AUGUST	1	0.06	
					ART	1	0.06	
					APARTMENT	2	0.12	
					ANIMALS	1	0.06	

これまでの助動詞、動詞、名詞をまとめると、7ヶ月あたりから徐々に語彙が増え、10ヶ月から一時期大量の語彙が産出され、また、同じ時期にエラー(不規則動詞に“-ED”をつけるエラー)も頻繁に現れることが特徴としていえるだろう。

3.4 General service list との比較

GSLの語彙リストとの比較は現在進行中であるが、現在のところ、3人のこどもの語彙全体とGSLを比較した結果、2284語のうち、782語が同じであることが分かった。これは全体の約34%になる。しかしながら、GSLの語彙リスト2274語は頻度順になっており、後半の1000語以降については、3人のこどもの年齢では日常使わない単語が多く含まれているので、それらを加味すると英語母語話者の一般語彙と小池コーパスの語彙のカバー率はより高くなる可能性がある。例えば、Sachikoの1-3ヶ月の語彙リストとの比較では、発話数116の中で、GSLとの共通語彙は、78個あり、そのうちGSLの最初の1000語に入るものが54個、残りの語彙リストに入るものが26個であった。全体では最初の3ヶ月の語彙の約68%が英語母語話者が日常使用する語彙を使っており、その内訳は、頻度の高い日常語彙が47%、頻度の低い難易度の高い語彙が約23%となる。

今後は、子ども達の各3ヶ月毎の語彙リストとGSLを比較し、それぞれの共通語彙の比率を計算し、3ヶ月毎の子どもの使用語彙が、母語話者の日常語彙とどの程度同じなのかを調べる予定である。しかしながら、前述のようにGSL語彙リストは、屈折語はまとめて一語にカウントするために、GSLを使用した語彙の分析では、文法的な発達を測るのは難しいといえる。これには、WordSmithなどの分析ツールをつかったさらに詳しい計算をする必要があると思われる。

3.5 英語学習者語彙リストとの比較

こちらの分析も進行中である。本来ならば、日本人の子どもが自然な状態で英語を習得するときの段階的な語彙リストがあれば良いのであるが、そのような語彙リストは存在しないので、それに近い語彙リストを探した結果、中学、高校などで使用する語彙リストを使用するのが現時点では妥当であると考えた。日本における英語語彙習得の研究は、そのほとんどが英語教育現場におけるものである。最近、コーパス言語学の発達により、語彙の習得に注目が向けられるようになってきた。これまでに作られた語彙リストとしては、「日本人1200人のスピーキングコーパス」(2004)、大学英語教育学会が作成した「JACET8000英単語」(2005)またその改訂版(2012)などが作られている。JACET8000の語彙リストは、レベル分けがされており、Level 1(順位1000位まで)は中学終了程度、Level 2(1001~2000)は高校終了程度と分けられている。そこで、JACET8000語彙のLevel 2までの2000語を小池コーパスと比較することにした。まずは、これらのJACET2000語が英語母語話者の語彙リストであるGSLと比較して、どの程度カバーしているのか(同じ語彙があるのか)についておおまかに把握することにした。比較の結果、2284語のうち、1475語がJACET2000語とGSL語彙リストの共通語彙であることがわかった。つまり、JACET2000語の語彙はGSL語彙の約64.46%をカバーしていることなる。この数値は高いとは言えないが、半数以上をカバーしている点では、英語母語話者の使用語彙とそれほど離れているとも言えない。今後の分析としては、JACET2000語彙とGSL語彙の共通語彙と小池コーパスデータ語彙とを比較したい。

3.6 まとめと考察

今回の分析はまた中途であるが、考慮する点が浮き彫りにされているので、今後の課題として明記する。

・語彙数の数え方の問題：

WordSmith では、異なり語はすべてひとつと数えているので、GSLのようなhead word または見出し語（レンマ）のような数え方をしない。しかし、今後の研究のためには、head word でカウントすることも必要なのではないか。

・エラー分析の可能性：

テキストにはエラータグがついているので、エラー分析が可能である。まだ手付かずの状態であるが、長期データにおける文法的形態素の分析は、時代を経ても意義のある研究となるのではないだろうか。

・WordSmithの機能を使ったその他の分析の可能性：

WordASmith にはコンコーダンス機能がついているので、特定の語についてのコロケーションなども分析できるのではないか。

今後の分析としては、文構造タグのついたデータファイルを集計し、3人のこどもの統語的な発達を調べる計画である。基本的にはPOS品詞タグ付与データ、文型タグデータを中心とした分析を継続して行うことになるが、今回の語彙の分析のようにテキストデータの活用も分析範囲とすることを考えている。

参考文献

- 和泉絵美・内元清貴・井佐原均（編）（2004）．『日本人 1200 人の英語スピーキングコーパス』 東京：アルク
- 小池生夫他（編集主幹）、寺内正典・木下耕児・成田真澄（編）（2004）『第二言語習得研究の現在：これからの外国語教育への視点』 東京：大修館
- 投野由紀夫（1997）『英語語彙習得論』東京：河原社
- 『大学英語教育学会基本語リスト JACET8000 List of Basic Words』（2003） 大学英語教育学会 基本語改定委員会編
- Bialystok, E. (1990). *Communication strategies: A psychological analysis of second-language use*. Oxford: Basil Blackwell.
- Brown, R. (1973). *A first language*. Cambridge: Harvard University Press.
- Koike, I. (1983). *Acquisition of grammatical structures and relevant verbal strategies in a second language*. Tokyo:Taisyukan.
- S. Gass & L. Selinker (2008). *Second Language Acquisition: An Introductory Course (3rd Edition)*. New York: Routledge/Taylor Francis.
- Quark, R, Greenbaum, S., Leech, G. and Svrtvik, J. (1972). *A grammar of contemporary English*. London: Longman.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Basingstoke: Palgrave Macmillan.
- Wells, G. (1985). *Language development in the pre-school years*. Cambridge: Cambridge University Press.

資料4 小池コーパスデータ情報

ファイル名	データ採集日	分析対象外	欠/未	Task	タスク詳細	POSタグファイル	作業状態	文型タグファイル	作業状態	備考
001	972/09/27			発音	絵の中に見えるものを英語で言う(単語)	POS001	済	文型_001	済	
002	1972/9/27			発音	日本語を英語にする	POS002	済	文型002	済	
003	1972/10/3			インタビュー;発音		POS003	済	文型003	済	
004	1972/10/5			インタビュー		POS004	済	文型004	済	
005a	不明	外		イミテーション;インタビュー	単語のリポート	POS005	外	文型005	済	
005b	不明	外		イミテーション	単語のリポート	POS005	外	文型005	未	作業不要
006	1972/10/8			イミテーション;インタビュー		POS006	済	文型006	済	
007	1972/10/23			インタビュー		POS007	済	文型007	済	
008	1972/10/28	外		暗唱	アメリカ国民としての宣誓文?を暗唱する	POS008	外	文型008	済	
009	1972.11.01	外		暗唱	アメリカ国民としての宣誓文?を暗唱する	POS009	外	文型009	済	
010	1972/11/12			イミテーション;インタビュー		POS010	済	文型010	済	イミテーションの部分はず
011	1972/11/24			イミテーション;インタビュー	音が似ている単語の聞き取り	POS011	済	文型011	済	イミテーションの部分はず
012	1972/11			インタビュー		POS012	済	文型012	済	
013	1972/12/7			インタビュー	JSN以外のこどもの発話の主(メモより)	POS013	済	文型013	済	
014	1972.12.12; 1973.01.05			プレイハウス		POS014	済	文型014	済	<head>が2つ
015	1972/12/29		欠(済)	インタビュー;歌	前半歌(Rudolf the red nose; Jingle Bell)	POS015	済	文型015	済	歌
016	不明			インタビュー		POS016	済	文型016	済	
017	1973/1/3		欠(済)	インタビュー		POS017	済	文型017	済	
018	1973/1/9			インタビュー		POS018	済	文型018	済	
019	1973.1.10/1.10	一部外	一部未(済)	インタビュー;英日訳	インタビューアーが絵本を読みそれを日本語に訳す。	POS019	済	文型019	済	<head>が3つ;日本語訳の部分のみはずす?
020	1973/1/12	外		英文和訳	インタビューアーの英語を日本語にする	POS020	外	文型020	済	こどもの発話は日本語
021	1973/1/13			インタビュー		POS021	済	文型021	済	
022	1973/1/16		一部未(済)	インタビュー		POS022	済	文型022	済	<head>が2つ
023	1973/1/20		一部未(済)	インタビュー		POS023	済	文型023	済	
024	1973/1/23		未(済)	インタビュー	絵本を見ながら?	POS024	済	文型024	済	
025	1973/2/7			会話	playing house	POS025	済	文型025	済	
026	1973/2/8			会話	playing house	POS026	済	文型026	済	
027	1973/2/9			会話	playing house	POS027	済	文型027	済	
028	1973.3.7(?)			和文英訳		POS028	済	文型028	済	<head>が2つ
029	1973/2/15			日記?	簡単な自己紹介	POS029	済	文型029	済	
030	1973/2/22			和文英訳		POS030	済	文型030	済	
031	1973/2/25			和文英訳		POS031	済	文型031	済	
032	1973/3/4			和文英訳		POS032	済	文型032	済	
033	1973/3/7			和文英訳		POS033	済	文型033	済	
034	1973/3/8	外		文法タスク(否定文)	インタビューアーの言う英語を否定文にする	POS034	済	文型034	済	
035	1973/3/8	外		文法タスク(複数)	インタビューアーの言う英文の中の名詞を複数にして英文を言い換える。	POS035	済	文型035	済	
036	1973/3/10			会話		POS036	済	文型036	済	
037	1973/3/11			会話	playing house	POS037	済	文型037	済	
038	1973/3/15			会話	playing house	POS038	済	文型038	済	
039	1973/3/15			会話		POS039	済	文型039	済	
040	1973/3/16			会話		POS040	済	文型040	済	
041	1973/3/20			ストーリー再話	A Turtle and a Rabbit	POS041	済	文型041	済	
042	1973/3/20			インタビュー	Talk with PK	POS042	済	文型042	済	
043	1973/4/2			インタビュー		POS043	済	文型043	済	
044	1973/4/3			会話		POS044	済	文型044	済	
045	1975/4/4		欠(済)	会話		POS045	済	文型045	済	
046	1975/4/8			音読	絵本か本を音読?	POS046	済	文型046	済	
047	1973/4/10			ストーリー再話	話を創作している?	POS047	済	文型047	済	
048	1973/4/26			ストーリー再話	話を創作している?	POS048	済	文型048	済	
049	1973/4/16			インタビュー		POS049	済	文型049	済	
050	1973/4/19			会話		POS050	済	文型050	済	
051	1973/4/20			音読	"reading & translation"のタグがついているが、絵本か本を音読しているだけのように見える。訳があるとすればどの	POS051	済	文型051	済	

052	1973/4/27		欠(済)	モノローグ		POS052	済	文型052	済	
053	1973/4/30		欠(済)	インタビュー		POS053	済	文型053	済	
054	1973/02			モノローグ		POS054	済	文型054	済	
055	1973/5/5			モノローグ		POS055	済	文型055	済	
056	1973/5/14			インタビュー	Talk with PK	POS056	済	文型056	済	
057	1973/5/24			インタビュー		POS057	済	文型057	済	task tag 追記
058	1973/6/11			会話		POS058	済	文型058	済	
059	1973/6/12			会話		POS059	済	文型059	済	話者不明 タグ多数
060	1973/6/14			インタビュー		POS060	済	文型060	済	
061	1973/6/21		欠(済)	インタビュー		POS061	済	文型061	済	
062	1973/6/25		欠(済)	会話		POS062	済	文型062	済	
063	1973/6/26			インタビュー		POS063	済	文型063	済	
064	1973/7/7			会話	playing house	POS064	済	文型064	済	
065	1973/7/8			会話	playing picture	POS065	済	文型065	済	
066	1973/7/14			会話	playing with a doll	POS066	済	文型066	済	
067	1973/7/15			会話	playing a game	POS067	済	文型067	済	
068	1973/7/19			ストーリー再話		POS068	済	文型068	済	
069	1973/7/25			会話	playing house	POS069	済	文型069	済	
070	1973/7/25		未(済)	会話	game of Life(?)	POS070	済	文型070	済	
071	1973/7/25		未(済)	会話	playing Life(?)	POS071	済	文型071	済	
072	1973/7/27			インタビュー		POS072	済	文型072	済	
073	1973/8/1			ストーリー再話	Sambo (「ちびくろサンボ」)	POS073	済	文型073	済	
074	1973/8/9			ストーリー再話	Ugly Duck; Sambo	POS074	済	文型074	済	
075	1973/8/14			会話		POS075	済	文型075	済	
076	1973/8/22			会話		POS076	済	文型076	済	
077	1973/8/22		欠(済)	会話		POS077	済	文型077	済	
078	1973/8/29		欠(済)	会話		POS078	済	文型078	済	
079	1973/8/30			ストーリー再話	A Big Turnip (「おおきなかぶ」)	POS079	済	文型079	済	<s></s>改 行なし
080	1973/9/6			ストーリー再話	Three Little Pigs (「3匹のこぶた」)	POS080	済	文型080	済	
081	1973/9/10			インタビュー		POS081	済	文型081	済	<s>と<名 前>の順が 反対
082	1973/9/17			インタビュー; 歌; スト リー再話	nursery rymes; Jack and Beanstock, Cinderella, Snow White;	POS082	済	文型082	済	
083	1973/9/18			会話		POS083	済	文型083	済	
084	1973/9/22			会話	playing house	POS084	済	文型084	済	
085	1973/9/22			会話	playing house	POS085	済	文型085	済	
086	1973/9/22			会話	playing house	POS086	済	文型086	済	
087	1973/9/24			会話		POS087	済	文型087	済	
088	1973/9/30			会話		POS088	済	文型088	済	
089	1973/9/30			会話		POS089	済	文型089	済	
090	1973.9.30			会話	playing a game	POS090	済	文型090	済	
091	1973/10/2		未(済)	ストーリー再話	「ピノキオ」	POS091	済	文型091	済	
092	1973/10/25			ストーリー再話	Dick and Betsy	POS092	済	文型092	済	
093	1973/11/5			ストーリー再話	Dick and Betsy	POS093	済	文型093	済	
094	1973/11/16	外		コメント	発音についてKoike のコメ ント	POS094	外	文型094	済	
095	1973/11/16			会話		POS095	済	文型095	済	
096	1973/11/27			ストーリー再話	「あかずきんちゃん」	POS096	済	文型096	済	
097	1973/11/27	外		コメント	発音についてKoike のコメ ント	POS097	外	文型097	済	
098	1973/10/6			インタビュー		POS098	済	文型098	済	
099	1973/12/2			会話		POS099	済	文型099	済	
100	1973/12/		一部未タ グ(済)	会話		POS100	済	文型100	済	
101	1973/12/10		一部未タ グ(済)	インタビュー		POS101	済	文型101	済	
102	1973/12/11			会話	playing games	POS102	済	文型102	済	
103	1973/12/12			描写	"English for Children"	POS103	済	文型103	済	
104	1973/12/18			描写	"English for Children"	POS104	済	文型104	済	
105	1974/1/3			描写		POS105	済	文型105	済	
106	1974/1/9			会話		POS106	済	文型106	済	
107	1974/1/10			ストーリー再話	「おおかみと七匹のこやぎ」	POS107	済	文型107	済	

108	1974/1/24			ストーリー再話	「3匹のこぶた」他	POS108	済	文型108	済	
109	1974/1/27			会話	playing a game	POS109	済	文型109	済	
110	1974/1/28			会話	playing a game	POS110	済	文型110	済	
111	1974/1/29			会話		POS111	済	文型111	済	
112	1974/2/3			会話		POS112	済	文型112	済	
113	1974/2/12			インタビュー		POS113	済	文型113	済	
114	1974/2/25			ストーリー再話	"The Wonderful Porridge Pot"	POS114	済	文型114	済	
115	1974/3/9			会話		POS115	済	文型115	済	
116	1974/3/14			会話	playing	POS116	済	文型116	済	
117	1974/4/17			インタビュー		POS117	済	文型117	済	
118	1974/4/29			ストーリー再話?	(おそらく)絵本を描写している	POS118	済	文型118	済	
119	1974/5/11			描写?	(おそらく)絵本を描写している	POS119	済	文型119	済	
120	1974/5/12			描写?		POS120	済	文型120	済	
121	1974/5/30			描写?		POS121	済	文型121	済	
122	1974/5/31			ストーリー再話	「ガリバー旅行記」	POS122	済	文型122	済	
123	1974/5/31			ストーリー再話	「ガリバー旅行記」 その他	POS123	済	文型123	済	
124	1974/6/23			会話		POS124	済	文型124	済	名前タグと <s> タグ が反対(修正済)
125	1974/6/24			会話	playing a game	POS125	済	文型125	済	
126	1974/7/21			描写		POS126	済	文型126	済	
127	1974/6/27			インタビュー		POS127	済	文型127	済	
128	1974/7/12			インタビュー		POS128	済	文型128	済	
129	1974/7/16			インタビュー		POS129	済	文型129	済	
130	1974/9/3			インタビュー		POS130	済	文型130	済	
131	1974/9/11			ストーリー再話	「シンデレラ」	POS131	済	文型131	済	
132	1974/9/30			描写		POS132	済	文型132	済	
133	1974/11/9			描写		POS133	済	文型133	済	
134	1975/1/12			会話		POS134	済	文型134	済	
135	1975/2/5			会話	play house	POS135	済	文型135	済	
136	1975/3/13			ストーリー再話	「ピノキオ」	POS136	済	文型136	済	
137	1975/3/4			ストーリー再話	「ピノキオ」	POS137	済	文型137	済	
138	1975/3/28			会話		POS138	済	文型138	済	
139	1975/4/10			会話		POS139	済	文型139	済	