

日本語 MLU（平均発話長）のガイドライン

— 自立語 MLU および形態素 MLU の計算法¹ —

宮田 Susanne

Guideline for Japanese MLU

— How to compute MLU_w and MLU_m —

Susanne MIYATA

This paper describes in detail the procedures for manual computation of MLU_m (Mean Length of Utterance in Morphemes) and MLU_w (Mean Length of Utterance in Words) for Japanese, as well as the application of the MLU program in the framework of CHILDES. The theoretical background is outlined, and the advantages and problems of MLU and its application to Japanese are pointed out.

Keywords : 文法発達, 平均発話長, MLU, 日本語

Grammar Acquisition, Mean Length of Utterance, MLU, Japanese

1. 平均発話長 (MLU) とは

平均発話長 (Mean Length of Utterances, MLU) は英語を獲得する子どもの文法発達の指標として Brown (1973) によって提案されている。この説の背景には、文法の発達による文構造の複雑化にともない、文の要素が増えるという事実がある。例えば言語獲得初期の *teddy jump* [クマさんジャンプ] という発話はその後 *the teddy is jumping* [クマさんがジャンプしている] に変わり、冠詞、判定詞、動詞の活用語尾などが付くことにより長くなる。

Brown (1973) は、形態素 (morpheme) を発話長計算の単位にすることを提唱している。すなわち独立した意味語のほか活用語尾も単位となる。名詞の複数形の *-s* や動詞の活用語尾の *-s*、*-ed*、*-ing* が 1 形態素として認められる。上記の例の場合、*teddy/jump* は 2 形態素、*the/ teddy/ is/ jump/ -ing* は 5 形態素になる。

1.1. 単語単位 MLU

Brown (1973) が提案した MLU では、形態素が計算の単位になっていたが、その後、形態素 MLU に対して批判の声が上がった。オランダ語 (Arlman-Rupp et al. 1976)、ドイツ語 (Park 1981)、アイルランド語 (Hickey 1991) に関して、形態素 MLU では子どもの能力が過評価されると指摘された。それらの言語では、語幹のみの動詞が存在しないため、子どもは最初から語尾の付いた語形を使わざるを得ない。英語の *go* [行く] は一形態素であるが、ドイツ語で同じ意味の *geh-en* は二形態素になり、*geh*-だけの形式では使用されないのである。そのような言語では MLU を単語単

¹ 本原稿の一部は Miyata, S. (Ed.) (2009). *Development of a developmental index of Japanese and its application to speech developmental disorders*. Report of the Grant-in-Aid for Scientific Research (B)(2006-2008) No. 18330141, Nagoya: Aichi Shukutoku University に掲載してある。

位で計算したほうがよいという結論に至っている（「Mean Length of Utterances in Words (MLUw)」）。

1.2. 日本語の MLU

日本語においても文法発達は文の長さによって反映される。膠着語として日本語は活用語尾（形態素）が累加されるので、語形は複雑になるにつれ、長くなる。形態素の数が語尾の数と一致し、ほとんどの場合、形態素の境界線ははっきりしている。例えば 食べ+させ+られ+た は語幹+使役+受け身+完了形に分けられる²。

さらに年齢につれ助詞や助動詞の使用頻度も増加する。小椋他（1997）は、12ヶ月から27ヶ月まで（3ヶ月毎10名ずつ）の幼児60名を対象とした文法発達調査の結果、助詞および活用語尾を含む助動詞の異なり語数（タイプ）と述べ語数（トークン）が増えていることを示している（図1）。日本語は膠着語として、活用語尾（形態素）が並べられるような形を取っている。つまり語形の複雑さが語形の長さとも直結している。また規則性が高く、不規則動詞は数個しかない³。このような高い規則性および構造の透明性が子どもの言語獲得を容易にし（Aksu-Koç & Slobin, 1985）、早期の過剰な使用も防いでいる。その結果、MLU 計算においてもアイルランド語などで見られる過評価の問題はおこらない（Miyata 1999）。

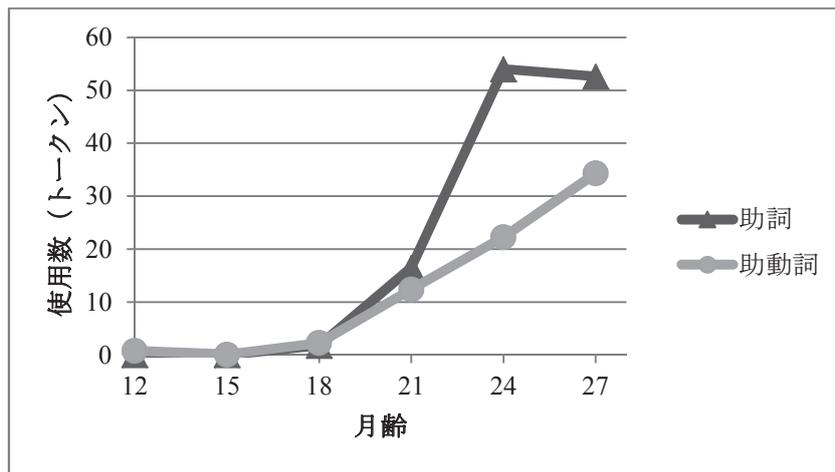


図1 12ヶ月～27ヶ月の助詞・助動詞の平均的な使用頻度の推移
(小椋他 1997:44 に基づく)

上記の形態素 MLU、自立語 MLU の議論を踏まえて、小椋他（1997）は「発話分割ガイド」（綿巻 1993、1994；綿巻 1999 付録 1、および綿巻・小椋 2004 にまとめてある）にもとづき、MLU を a) 「接辞・活用形付属語 MLU」、b) 「活用形付属語 MLU」、c) 「自立語付属語 MLU」および d) 「自立語 MLU」の 4 通りで計算した（表 1、図 2）。29 分の親子会話において、12 ヶ月で平均的に 10.4 の明瞭な発話、18 ヶ月で 56.8 発話、24 ヶ月で 162.5 発話を得た。明瞭な発話がなかった場合は MLU を 0 とした。その結果、特に「接辞・活用形付属語 MLU」と「活用形付属語 MLU」は、機能語（助詞、助動詞、付属語）の異なり語数および明瞭発話数と、高い相関を示した（それぞれ

² 以下の文法用語は増岡・田窪（1995）に基づく。なお、「イ形容詞」「ナ形容詞」は「形容詞」「形容名詞」と名付ける。
³ 不規則的な活用を示す動詞としてある、する、くる、いく（完了形のみ）、そしてござる、いらっしゃる、なさる、くださるのグループおよび判定詞だ³が挙げられる。そのほかに化石化した語形のちょうだい、おいで。

0.8 以上)。それに対し、「自立語付属語 MLU」および「自立語 MLU」は $r = 0.6$ 台であった。この研究の対象となった幼児は全員まだ MLU2.0 以下であり、Brown (1973) の段階 I にあたるので、より年齢の高い子どもの調査が必要である (小椋他 1997) が、低い年齢において形態素 MLU も有効であると考えられる。

表 1 日本語における MLU 計算法の種類

	自由語	助詞	活用形	接頭辞・接尾辞
ワード単位 MLU (綿巻 1999)				
自立語 MLU (小椋 1997)	○	×	×	×
自立語 MLU (MLUw; 宮田 1999)				
自立語付属語 MLU (小椋 1997)	○	○	×	×
活用形付属語 MLU (小椋 1997)	○	○	○	×
接辞・活用形付属語 MLU (小椋 1997)				
形態素単位 MLU (綿巻 1999)	○	○	○	○
形態素 MLU (MLUm; 宮田 1999)				

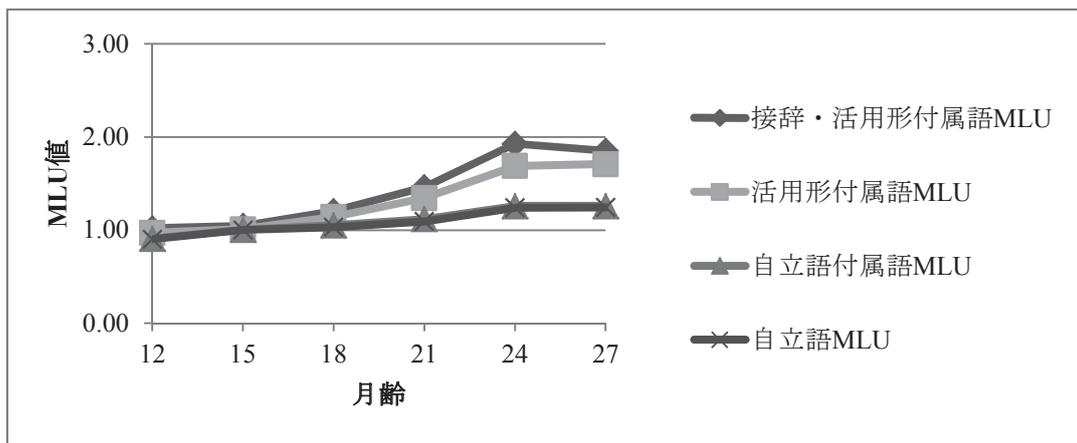


図 2 12ヶ月～27ヶ月の4種類の MLU の推移 (小椋他 1997:43 に基づく)

しかし、以上のような議論の際、忘れてはいけないことは、4通りの計算法のどれも欧米の MLUw (単語単位 MLU) と異なるものだという点である。ドイツ語、オランダ語、アイルランド語は英語と同様にアルファベット文字で書かれ、普段から単語に分けて書く言語である。その単語をそのまま MLU 単位として使用する発想であり、確かに最も身近な方法であろう。だが、自立語に限らず、文法語の一部も単語として分けて書かれる。例えば冠詞 (*the*)、前置詞 (*in, on*)、助動詞 (*will*)、否定詞 (*not*) などは MLUw の単位になる。つまり上記の日本語の自立語 MLU よりは文法語が多いことになる。

綿巻 (1999) は一人の子ども (F 児) の統合意味構造と助詞助動詞の発達を背景に「ワード単位 MLU」と「形態素単位 MLU」を解析し、MLU の成長が特に助詞の発達に影響されることを指摘した。3人の子どもの縦断的に調べた宮田 (1999) は形態素 MLU (以下 MLUm) と自立語 MLU (以下 MLUw; 小椋他の「接辞・活用形付属語 MLU」と「自立語付属語 MLU」に相当する) を計算し、3才までの推移を報告した。MLUm は初期に一時的に高い値が見られ、過評価している可能性を指摘している。図 3・4 は綿巻 (1999) の F 児、宮田 (1999) の Aki, Ryo, Tai, そ

して Tar (国立国語研究所 1982, 1983) の MLU の結果をまとめたものである。子どもの間にはかなりの発達のずれが見られるが、全員の MLU が徐々に上がることが分かる。MLUm の場合は 4.0 を超えた後は MLU 値が安定しなくなり、上がった、下がったりする。MLUm の平均値を見ると、2才半(30ヶ月)あたりに MLUm 3.0 を超え、3才近くなると MLUm 4.0 まで上がる(図3)。片方、MLUw は特に初期に安定しており、その段階ではよりよく表していると思われるが、その後、MLUw の増加の幅が小さいので、むしろ MLUm のほうが文法発達を敏感に捉えていると考えられる(図4)。この点では小椋他(1997)の結果と一致している。

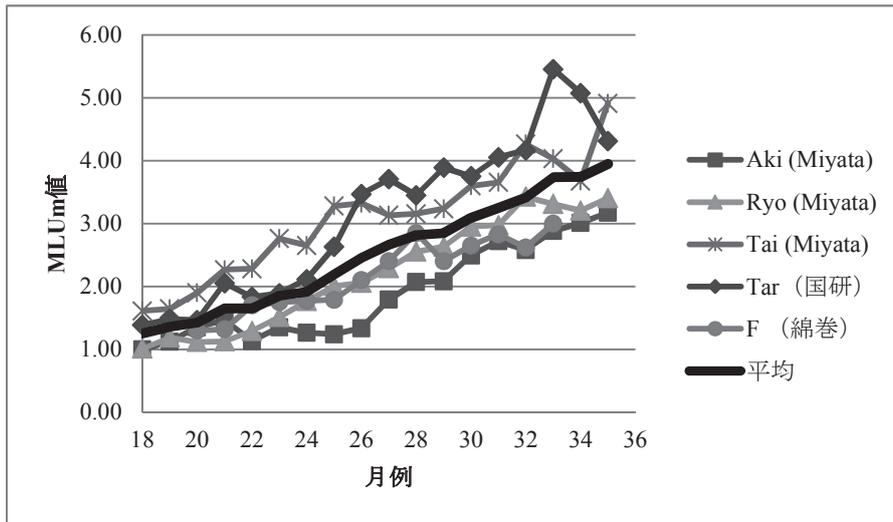


図3 5名の子ども(1才6ヶ月~3才0ヶ月)の形態素 MLU (MLUm) の推移

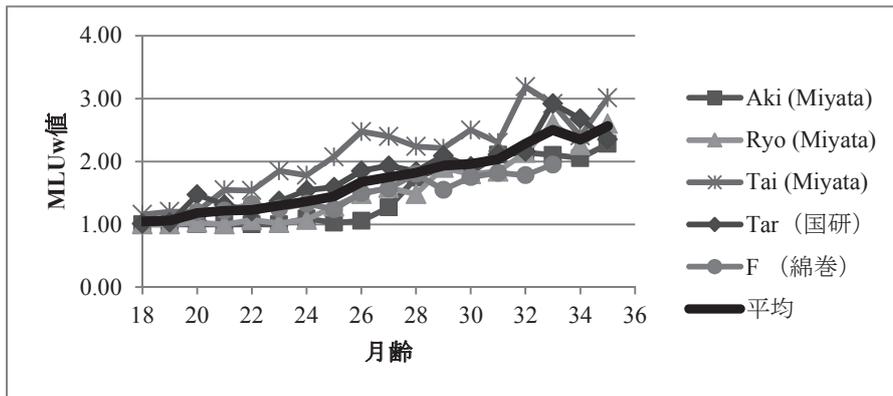


図4 5名の子ども(1才6ヶ月~3才0ヶ月)の自立語 MLU (MLUw) の推移

宮田(1999)は人数も3人に限られ、小椋他(1997)のように文法発達を表す助詞や助動詞の使用状況と比較していないため、それぞれの MLU 段階が表す文法能力もはっきりしない。今のところ言えることは、実際に MLU を使用する場合、初期段階(健常児の場合は2才~2才半まで)では MLUw (自立語 MLU) を計算し、その値が 1.5 を超えた時点で、MLUm (形態素 MLU) も計算することが望ましいということだろう。すなわち、文法がすでにある程度進んだ段階(健常児の場合2才半以上)では MLUm を使用することが推奨される。どのような計算法が日本語に一番合った MLU であるかは今後の研究課題である。

1.3. MLU 段階 (MLU stages)

Brown (1973) では MLU を 0.5 刻みで発達段階 (stages、表 2) に分け、それぞれの段階の特徴を述べている (表 1)。Chabon, Kent-Udolf & Egolf (1982) が示したように stage V (MLUm 4.0 以上) では MLU 値が不安定になり、文法発達指標として不適切である。

日本語の MLU 段階として Barner, Guerriero & Oshima-Takane (1999) の提案がある。Barner らが指摘しているように、MLU 段階に分けることは横断研究、および子どもの発達段階の比較において不可欠である。しかし、それぞれの MLU 段階の発達の特徴についてさらに確認する必要がある (表 3、表 4)。

表 2 英語の MLU stages (Brown, 1973 に基づく)

略称	段階	MLU 範囲	発達の特徴
I early	Early Stage I	1.0 - 1.49	意味役割を表す
I late	Late Stage I	1.5 - 1.99	"
II	Stage II	2.0 - 2.49	意味の変調：基本的な 14 つの文法的形態素の獲得
III	Stage III	2.5 - 2.99	単文の変調：助動詞の獲得
IV	Stage IV	3.0 - 3.99	複文構造
V	Stage V	4.0 以上	

**表 3 日本語の MLU 段階：接辞・活用形付属語 MLU (MLUm) の場合
 (Barner 他, 1999 に基づく)**

略称	段階	MLUm 範囲
EI	Early Stage I	1.00 - 1.49
LI	Late Stage I	1.50 - 1.99
II	Stage II	2.00 - 2.49
III	Stage III	2.50 - 2.99
EIV	Early Stage IV	3.00 - 3.49
LIV-EV	Late Stage IV - Early Stage V	3.50 - 3.99
LV	Late Stage V	4.00 - 4.49
PV	Post Stage V	4.50 以上

表 4 日本語の MLU 段階：自立語 MLU (MLUw) の場合 (Barner 他, 1999 に基づく)

略称	段階	MLUw 範囲
wI	Stage I	1.00 - 1.49
wII	Stage II	1.50 - 1.99
wIII	Stage III	2.00 - 2.49
wIV	Stage IV	2.50 - 2.99
wV	Stage V	3.00 - 3.49
wPV	Post Stage V	3.50 以上

2. MLU 計算の準備

2.1 発話の切り方

「子どもの発話を書き起こす際、一つの発話として区切るべきかどうかは、発話間の休止、およびイントネーションから判断して決める」(Bloom & Lahey, 1978) ことが原則である。しかしそのような原則があっても実際には区切るべきか否かで迷うことがある。その際は、MLU は文法発達を計る指標なので、「文」のつながりを意識すること必要である。文法的につながっていないものは、停止がなくても区切るべきである。

子: ほら見てみんな入った!

⇒ 発話1: ほら見て!

⇒ 発話2: みんな入った!

逆に、以下の例のように、相手の相づちが入った場合、相づちによって切られてはいるが、「これをここに置いて出来上がり」のように一つの発話として扱いたい。

子: これを ここに入れて 出来上がり!

母: うん。 うん。

また、倒置も文の一部として考える。その際、倒置された部分が低いピッチになることが判断の手がかりになる。

入れてね、早く。 < 早く入れてね。

2.2. MLU 計算から省くもの

解析対象は、連続した発話であるが、以下の発話(部分)は対象外として省く。

2.2.1. 不明瞭な発話

途中で放置された発話、不明瞭な発話（不明瞭な部分を含む発話も含む）、および聞き取りに自信がもてない発話は解析対象から除外する。

豆もおいてみ…あ、落ちた! → 前半の発話を省く

チョコ [不明な部分] だから。 → 発話全体を省く

えびせん [?] ちょうだい。 → 発話全体を省く（「えびせん」を言ったかどうかは不確定）

2.2.2. 繰り返し

繰り返し部分は省く。発話全体が繰り返された場合は、1 回目の発話のみを解析する。発話の一部が繰り返された場合は、その繰り返しの部分だけを省く。ただし、繰り返しの一部分においても違いが見られる場合は、省かずに解析対象とする。また何回も言いかけている単語は一つとして数える（吃音症など）。

みかん落ちた。みかん落ちた。 → 2 番目の発話を省く

チョコ欲しい、欲しい、欲しい。 → チョコ欲しい として解析する

ぼくも、ぼくも、ぼくも行く! → ぼくも行く として解析する

熊さんおちる。熊さんおちる! → 両発話を解析する

食べる? 食べる。 → 両発話を解析する

た、た、た、たべた。 → たべた として解析する

2.2.3. 挨拶表現

こんにちは、バイバイ、おはよう、いただきます、お帰りなさい のような挨拶表現は省く。

しかし、文の一部として使われた場合は解析の対象にする。

プーさん、こんにちは。 -> 発話全体を省く
先生にこんにちはゆった。 -> こんにちは を 1 としてカウント

2.2.4. 相づち、間投詞、フィラー、否定、肯定の用言

はい、いいえ、ええ、うん、ううん、ああ、ええと、あのう、イエー などのフィラーや相づちは、解析対象外として省く。

ええとどこ行った? -> どこ行った として解析する

2.2.5. 歌詞など丸暗記していると思われる文章

歌詞、絵本の文章、CM など、子どもが丸暗記していると思われる発話は、子どもによって自ら作られた文構造ではないため、解析対象から省く。連続で数えることやエービーシーを暗唱することも丸暗記に含まれる。

咲いた、咲いた、チューリップが。 -> 発話から省く
いい味ヤマモリ -> 発話から省く
いち、にい、さん、よん、ごお、ろく、なな! -> 発話から省く

2.2.6. 擬音語・擬態語

ガーガー、キーキー、ころころころのような擬音語、擬態語は省く。しかし、名詞または副詞として文の中で使われた場合は解析の対象にする。

けろけろ。 -> 発話を省く
けろけろちゃんが出た。 -> けろけろちゃん を名詞として解析に含む
ぶうぶうが来た。 -> ぶうぶう を名詞として解析に含む
ゲラゲラ笑った。 -> げらげら を副詞として解析に含む

2.3. 発話数

多くの研究者は100発話単位でMLUを計算している(排除すべき発話はこの数字にすでに含まれていない)。また、Brown (1973) は最初の25発話はMLU計算から除外すべきであるとしている。子どもが落ち着いて、会話に集中しているデータに限るべきであるという理由からであろう。しかしながら初期段階で100発話を集めることは困難であることも多い。Rondal & DeFays (1978) は40名の幼児(1;8~2;8)の全発話を、25、50、75、100、125、150、175、200発話サンプルに切り、MLU計算を行った結果、50発話以上のサンプルではMLUの信頼性がほとんど変わらなかった。Miyata (1999) では75発話サンプル、100発話サンプル、150発話サンプルのいずれについても有意な相関が認められた(MLU_w、MLU_mともに $p < .01$ 水準)。しかし初期を超えた段階(MLU_m3.0以上、またはMLU_w2.0以上)では有意な相関が認められるものの、下がる傾向が示された。

初期段階では50発話で充分である可能性が高いが、助詞や活用が多くなっていく段階では100発話のサンプルが望ましいというのが、今のところの結論であろう。

3. MLU_w (自立語 MLU) の計算法

本節で紹介する自立語 MLU (以下 MLU_w) の計算法のガイドラインは、綿巻 (1999) および小椋他 (1997) と一致している。MLU_w では接頭辞、語尾、助詞などの付属語は、個別にカウントされない。

3.1. 名詞・固有名詞

名詞および固有名詞は付属語を含めて1語として計算する(表5、例1~11参照)。名字と名前は別の単語として分ける(例6)。

3.2. 助詞

格助詞、終助詞、取立助詞、接続助詞はカウントされない(例12~18)。

格助詞：から、が、で、と、に、の、へ、まで、を、より

終助詞：か、かい、かな、つけ、さ、ぞ、なあ、ね、ねん、の、もん、わ、や、よ 等

取立助詞：こそ、くらい、しか、とか、として、とって、とも、のみ、など、ほど 等

接続助詞：が、か、から、けど/けれど/けれども、し、と、のに、ながら、も 等

しかし発話の最初に使われる接続詞は1語として数える(例19~22)。

接続詞：しかし、そして、だから、もし、それなら/ほんなら/ほな、それで/そいで/ほんで、
それでは/それじゃ/ほんじゃ、だけど、で、でも 等

3.3. 動詞、形容詞、助動詞、判定詞

動詞の活用語尾は個別にカウントしない(例23~46)。~(て)あげる、~(て)みるのような助動詞もカウントしない(例49~53)。判定詞(だ、ですなど)もカウントしない(例61~67)。形容動詞の場合も、判定詞(だ、な、で、に、です、だった等)の部分はカウントしない(例69~70)。形容詞の活用語尾も同様に個別に数えない(例54~60)。しかし軽動詞のする、なるは独立した動詞として数える(例47~48、71~72、82)。動作名詞との組み合わせでもするは独立した軽動詞としてカウントされる(例73~75)。

3.4. 複合語および数字

複合語は分けない(例76~79)。数字は百、千、万単位で分ける(例80~84)。助数詞は分け
ない(例85~91)。

4. MLUm(形態素MLU)の計算法

MLUm(自立語MLU、または接辞・活用形付属語MLU)では「子ども+たち+が」語幹+語尾+格助詞や「食べ+られ+た」語幹+受け身+完了形(両方3形態素)というように単語と付属語を分けて計算を行う。以下で、具体的な分け方を示し、よく現れる語尾などを品詞別にまとめる。形態素を「+」で分け、隣に形態素数を記す。形態素の分割が未完了形の扱い以外、綿巻(1999、付録1)とほぼ一致する。⁴

4.1. 名詞・固有名詞

名詞の場合は語幹のほかに語尾や接頭辞をそれぞれ別個に数える(例1~9)。

名詞に付く接頭辞の例：お~(御+魚)、こ~(子牛、小牛)、ご~(御本)、だい~(大失敗)、だい~(第三)、ふ~、ぶ~(不)、ひ~(非+現実的)、まい~(毎朝)、りょう~(両足)

名詞に付く語尾の例：~くん(カバ君)、~っこ(端っこ)、~ごっこ(赤ちゃん+ごっこ)、~さん(ねずみ+さん)、~たち(子ども+達)、~ちゃん(猫+ちゃん)、~だらけ(粘土+だらけ)、~ちゅう、~じ

⁴厳密さを欠く場合もあるが、分かりやすさを優先し、ひらがな表記を用いる。例えば、食べますは厳密にはtabe-mas-uというようなわけ方になるが、食べ+ま+すと表記する。なお、文法の解説および用語は益岡・田窪(1995)に基づく。

表 5 自立語 MLU および形態素数 MLU の分け方と単位数の例一覧

例	自立語 MLU の分け方	自立語数	形態素 MLU の分け方	形態素数
名詞・固有名詞				
1	お客様	1	お+客+さま	3
2	あきふみくん	1	あきふみ+くん	2
3	石塚先生	1	石塚+先生	2
4	田中選手	1	田中+選手	2
5	レイナちゃん	1	レイナ+ちゃん (他人の場合)	2
6	田中+ミツちゃん	2	田中+ミツ+ちゃん (他人の場合)	3
7	ナオキ君	1	ナオキ君 (本人の場合)	1
8	お母さん	1	お母さん (名前代わりに使われている場合)	1
9	お母さん	1	お+母+さん (名詞として使われている場合)	3
10	ミッキーマウス	1	ミッキーマウス	1
11	ウルトラマン	1	ウルトラマン	1
助詞				
12	どんぐりが	1	どんぐり+が	2
13	箱には...	1	箱+に+は...	3
14	お茶でも+どうぞ。	1	お茶+で+も+どうぞ。	4
15	リンゴも+落ちたよ	2	リンゴ+も+落ち+た+よ。	5
16	食べたよ。	1	食べ+た+よ。	3
17	おいしいから+たべてね!	2	おいし+い+から+食べ+て+ね!	6
18	動物だから。	1	動物+だ+から。	3
19	だから+行かない。	2	だから+行+か+ない。	4
20	だって+ほしいもん。	2	だって+ほし+い+もん。	4
21	それじゃ+行くか?	2	それじゃ+行+く+か?	4
22	それから+どうした?	3	それから+どう+し+た?	4
動詞				
23	食べる	1	食べ+る	2
24	食べた	1	食べ+た	2
25	食べて	1	食べ+て	2
26	食べよう	1	食べ+よう	2
27	食べる	1	食べ+る	2
28	食べれば	1	食べ+れば	2
29	食べたら	1	食べ+たら	2
30	食べたり	1	食べ+たり	2
31	食べなさい	1	食べ+なさい	2
32	食べられた	1	食べ+られ+た	3
33	食べさせた	1	食べ+させ+た	3
34	食べちゃった	1	食べ+ちゃ+った	3
35	食べれた	1	食べ+れ+た	3
36	食べはる	1	食べ+は+る	3

表5 (続き)

例	自立語 MLU の分け方	自立語数	形態素 MLU の分け方	形態素数
37	食べます	1	食べ+ま+す	3
38	食べない	1	食べ+な+い	3
39	食べん	1	食べ+ん	2
40	食べなくて	1	食べ+なく+て	3
41	食べなかった	1	食べ+な+かつ+た	3
42	食べないで (命令形)	1	食べ+ない+で	3
43	食べなくちゃ	1	食べ+なく+ちや	3
44	食べたい	1	食べ+た+い	3
45	食べたかった	1	食べ+た+かつ+た	3
46	食べさせられなかった	1	食べ+させ+られ+な+かつ+た	5
47	お帰りに+なる	2	お+帰+り+に+な+る	6
48	食べなく+なる	2	食べ+な+く+な+る	5
助動詞				
49	食べてみる	1	食べ+て+み+る	4
50	食べて(いる)	1	食べ+て+い+る	4
51	食べてあげる	1	食べ+て+あ+げ+る	4
52	食べて下さい	1	食べ+て+下+さい	3
53	食べてごらん	1	食べ+て+ご+らん	3
形容詞				
54	美味しい	1	美味+し+い	2
55	美味しく	1	美味+し+く	2
56	美味しかった	1	美味+し+か+つ+た	3
57	美味しくない	1	美味+し+く+な+い	3
58	美味しくなかった	1	美味+し+く+な+か+つ+た	3
59	粘土臭い	1	粘土+臭+い	3
60	子どもっぽかった	1	子ども+っ+ぽ+か+つ+た	3
判定詞				
61	パパだ	1	パパ+だ	2
62	パパだった	1	パパ+だ+つ+た	3
63	パパじゃない	1	パパ+じや+な+い	4
64	パパです	1	パパ+で+す	3
65	パパでした	1	パパ+で+し+た	3
66	パパでしょう	1	パパ+で+し+よ+う	3
67	美味しいんだ	1	美味+し+い+ん+だ	4
68	美味しかったでしょう (!)	1	美味+し+か+つ+た+で+し+よ+う (!)	3
形容名詞				
69	きれいな+人	2	きれい+な+人	3
70	きれいです	1	きれい+で+す	3
71	きれいに+なった	3	きれい+に+な+つ+た	4
72	きれいに+する	2	きれい+に+す+る	4

表5 (続き)

例	自立語 MLU の分け方	自立語数	形態素 MLU の分け方	形態素数
動作名詞				
73	勉強+する	2	勉強+す+る	3
74	ジャンプ+する	2	ジャンプ+する	3
75	勉強+しない	2	勉強+し+な+い	4
複合語				
76	バナナジュース	1	バナナ+ジュース	2
77	消防自動車	1	消防+自動車	2
78	何ジュース	1	何+ジュース	2
79	食べはじめる	1	食べ+はじめ+る	3
数字・助数詞				
80	じゅういち (11)	1	じゅう+いち	2
81	さんじゅう (30)	1	さん+じゅう	2
82	さんじゅうご (35)	1	さん+じゅう+ご	3
83	にひゃく+じゅう (210)	2	に+ひゃく+じゅう	3
84	さんまん+きゅう+せん+ はっぴゃく+ごじゅう+えん	4	さん+まん+きゅう+せん+はっ+ひゃく+ ご+じゅう+えん (39,850 円)	9
85	三枚	1	三+枚	2
86	ふたり	1	ふた+り	2
87	一番に+なった	2	一+番+に+なっ+た (順番を表す場合)	5
88	一番+高い	2	一+番+高+い (副詞的な使い方の場合)	3
89	もう+一回	2	もう+一+回 (幼児語として例外)	2
90	もう+一つ	2	もう+一+つ (幼児語として例外)	2
91	ひとり+で	1	ひとり+で (幼児語として例外)	2

ゅう (休み+中)、～じん (アメリカ+人)、～つき (ご飯+付き)、～ばん (3+番)、～ばんめ (3+番+目)、～や (果物+屋)、～やさん (魚+屋+さん)、～ゆき (東京+行き)、～よう (子ども+用)

以下の紛らわしい単語は1形態素として扱う。

- 子ども 1 語幹「子」は単独でも使われるが、「ども」の複数型がほかに使われないので、1形態素として扱うが、お+子+様 は接頭辞+語幹+語尾として分ける。
- お茶 1 語幹「ちゃ」だけは使わないので「お茶」を1形態素として扱う。
- ご飯 1 語幹「はん」だけは使わないので「ご飯」を1形態素として扱う。
- あかちゃん 1 語幹「あか」は別の意味になるほか、「あかさま」などの変化は使われない。

固有名詞には1形態素としてカウントされる(例2~7、10~11)が、名字と名前は別単語として分ける(例6)。～さん、～ちゃん、～先生などの語尾が付く場合は、原則として語幹と語尾と別個にカウントされる(例2~6)。

しかし、初期発達段階で固定された形で使われている子供本人、および家族の名前は分けない(例7)。同じく、お母さん、お姉さんなどの親族を表す名詞も、固定された形で名前代わりに使われている場合は一形態素として数える(例8)。2つ以上の単語からできたキャラクター名でも一形態素として数える(例10~11)。固定された使い方であるかどうかは、観察者の判断で決めてよい。それに対して友達の名前の場合は、分ける。

4.2. 助詞

助詞のが、を、の、は、等の助詞は別個に数える（例 12～16）。

格助詞： から、が、で、と、に、の、へ、まで、を、より

提題助詞： は、も、って

取立助詞： こぞ、くらい、しか、とか、として、とって、とも、のみ、など、ほど 等

引用助詞： と、って

終助詞： か、かい、かなあ、つけ、さ、ぞ、なあ、ね、ねん、の、もん、わ、や、よ 等

文と文を繋ぐ接続助詞は一形態素としてカウントする。（例 17～18；動詞、形容詞の活用形の分け方については 4.3 節を参照。）

接続助詞： が、か、から、けど／けれど／けれども、し、と、のに、ながら、も 等

接続助詞に対し、接続詞は発話の最初に使われる。接続詞は、判定詞の「だ～」や「それ～」で始まるものが多いが、それらは一かたまりの一形態素として扱われる（例 19～22）。

接続詞： しかし、そして、それなら／ほんなら／ほな、それで／そいで／ほんで、だから、だけど、もし、
それでは／それじゃ／ほんじゃ、それから、で、でも、 等

4.3. 動詞・形容詞・助動詞・判定詞

動詞の語尾は、別個にカウントする（例 23～31）。時制の語尾の前に受け身、使役などの語尾が入っている場合は、それらも別個に数える（例 32～46）。

ほかの動詞との組合せの場合は、動詞ごとで分け（例 47～53、79）、さらに～てで接続されている場合は～ても別個にカウントする（例 49～53）。また、～下さい、～ちょうだい、～ごらんは一形態素としてカウントする（例 52～53）。

形容詞も同じく活用語尾が個別に数えられる（例 54～58）。名詞を形容詞に変える語尾に活用形が付いた場合は、活用形も分けてカウントする（例 59～60）。

判定詞（コンピュータ）は基本的に動詞と同様に扱うが、不規則な活用のため MLU 計算に影響する（例 61～66）。のだ（んだ）も助詞 の と判定詞 だ として別個に数える（例 67）。また、です、でしょう、じゃん は時制の意味を表さず、一かたまりとして終助詞のように使われている活用形もある（例 68）。

動作名詞（サ変動詞）と形容名詞（形容動詞、またはナ形容詞）は名詞の部分と軽動詞や助詞の部分に分ける（例 69～75）。

4.4. 複合語・数字

複合語は単語毎に分ける。複合語は名詞＋名詞（例 76～77）に限らず、疑問詞を含む複合語（例 78）と複合動詞（例 79）のように様々な組み合わせがある。

数字（11 以上）も複合の一種として扱う（例 80～84）。助数詞（例：～枚、～冊）は一形態素として数える（例 85～87）。しかし、一番も副詞的な使い方の場合、1 形態素として扱う（例 88）。また、幼児言語で他の助数詞より遥かに早く現れる（もう）いっぱい、（もう）ひとつ および一人（で）はそれらを一かたまりの一形態素としてカウントする（例 89～91）。

5. CHILDES で MLU の自動計算を行う場合

発話データベース CHILDES の付属プログラム CLAN を用いて MLU_w、MLU_m を自動的に計算することができる。CHILDES に関する基礎的情報は <http://www2.aasa.ac.jp/people/smiyata/CHILDESintroJPN.pdf> にまとめられている。CHILDES 全体のホームページ<<http://childes.psy>

cmu.edu/> から発話データおよび CLAN プログラムを無料でダウンロードすることができる。入門書としては宮田・村木・森川 (2004) がある。

CHILDES で公開されているデータの一部では、すでに形態素タグが付いている。以下の例のように*CHI の発話行の下に%mor のティアで形態素がハイフン (*kuma-san*、*ochi-PAST*) で分けてある。その情報に基づいて形態素数は算出される。

```
*CHI:   kumasan ga ochita .  
%ort:   熊さんが落ちた。  
%mor:   n|kuma-san ptl:case|ga v:v|ochi-PAST .
```

上記のような形になっている場合は次のコマンドで自立語 MLU または形態素 MLU 計算できる。ファイルのチェックおよびコマンドボックスの呼び出し方などについては宮田他 (2004:48ff) を参照。コマンドはプログラム (*mlu*)、オプション (+/-f, b, s, t, z など) およびファイル名 (**.cha*) の3部からできている。今回使用するオプションは -f (ファイルに保存せず、画面出力する)、+/-b (形態素として認めるまたは認めない複合語[b+], 接頭辞[b#], 接尾辞[b-])、-s (解析から外す品詞)、+z (解析範囲: 1~100 番目の発話)、+t (話者コード) である。

MLUw (自立語 MLU) の CLAN コマンド:

```
mlu -f -b+ -b# -b- -s"[+ *]" -s"ptl:*" -s"v:cop*" -s"co*" -s"onoma|*"
+z1u-100u +t*CHI *.cha
```

MLUm (形態素 MLU) の CLAN コマンド:

```
mlu -f +b+ -s"[+ *]" -s"co*" -s"onoma|*" +z1u-100u +t*CHI *.cha
```

アウトプットとして発話数 (number of utterances)、形態素数 (number of morphemes)、MLU (ratio of morphemes over utterances) とその標準偏差 (standard deviation) が出力される。

%mor ティアが付いていない場合は、a) CHAT フォーマットである、b) ローマ字表記であり、Wakachi2002 (宮田 2003) の分かち書きに従っている、という2つの条件を満たせば、MOR コマンドで%mor ティアを自動的に付けることが可能である。

宮田他 (2004) の第4章に簡単にまとめてある。図5のようにファイルの最初と最後に@Begin および@End を入力し、使用言語 (Languages)、参加者 (Participants)、個人情報 (@ID) を加える。発話は話者記号 (*CHI など) の後に、単語 (名詞、動詞、助動詞、助詞など) で分けてヘボン式で入力する。

分かち書きガイドラインは <http://chilides.psy.cmu.edu/morgrams/wakachi2002.zip> でダウンロードできる。CHILDES フォーマットのファイルに MOR コマンドで%mor ティアを付けることができるが、その使い方については宮田他 (2004) 第7章、および <http://www2.aasa.ac.jp/people/smiyata/JMORnoTsukaikata.pdf> で説明してある。日本語の MOR に必要なライブラリーファイルは <http://chilides.psy.cmu.edu/morgrams/jpn.zip> に置いてある。

6. 最後に

MLU が日本語に関しても有効な発達指標として認めて来たが、さらに研究が必要であろう。自立語 MLU (MLUw) が何才まで有効であるか、形態素 MLU (MLUm) の初期から使えるか、ま

た MLU 段階の発達的な特徴を持っているのかについてさらに調べる必要がある。複数の子どもの縦断的データを解析し、標準化を行う必要がある。さらにサンプルサイズに関する研究も望ましく感じる。初期は 50 発話でも充分であることがあり得るが (Miyata, 1999)、3 才を超えた時に何発話で信頼性のある MLU 値が得られるのかを調べる必要がある。

図5 CHILDESフォーマットの例

```
@Begin
@Languages:      jpn
@Participants:  CHI Child, MOT Mother
@ID:             jpn||CHI|3;0.0|male|||Target_Child||
@ID:             jpn||MOT|||||Mother||
*CHI:           kore nani ?
*MOT:           mikan deshoo .
*CHI:           mikan tabetai yo .
@End
```

文献

- Aksu-Koç, A.A. & Slobin, D.I. (1985). The acquisition of Turkish. In D.I.Slobin (ed.) *The crosslinguistic study of language acquisition. Vol.1*, 839-78.
- Arlman-Rupp, A.J.L., van Niekerc de Haan, D., & van de Sandt-Koendermann, M. (1976). Brown's early stages: some evidence from Dutch. *Journal of Child Language* 5. 267-274.
- Barner, D., Guerriero, A.M.S., & Oshima-Takane, Y. (1999). The calculation of Mean Length of Utterance: Methods for assigning transcripts to Brown's linguistic stages. Unpubl. Paper. McGill University.
- Bloom, L. & Lahey, M. (1978). *Language development and language disorders*. John Wiley:New York.
- Brown, R. (1973). *A first language*. Cambridge, Mass.: Harvard University Press.
- Chabon, S.S., Kent-Udolf, L., & Egolf, D.B. (1982). The temporal reliability of Brown's Mean Length of Utterance (MLU-M) Measure with post-stage V children". *Journal of Speech and Hearing Research* 25. 117-124.
- Crystal, D. (1974). Review of R. Brown (1973), *A First Language*. *Journal of Child Language* 1. 289-307.
- Hickey, T. (1991). Mean length of utterance and the acquisition of Irish. *Journal of Child Language* 18. 553-569.
- Klee, T. & Fitzgerald, M.D. (1985). The relation between grammatical development and mean length of utterance in morphemes. *Journal of Child Language* 12. 251-69..
- 国立国語研究所 (1982-1983) . 『幼児の言葉資料(3～6). 2～4 歳児の言葉記録』 秀英出版.
- 益岡隆志・田窪行則 (1995) . 『基礎日本語文法 改訂版』 くろしお出版.
- MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk. Third Edition*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Miller, J.F., & Chapman, R.S. (1981). The relation between age and mean length of utterance in morphemes. *Journal of Speech and Hearing Research* 24. 154-161.

- Miyata, S. (1999). Assigning MLU stages for Japanese. *Journal of Educational Systems and Technologies. The Audio Visual Center, Chukyo University Nagoya Japan* 9. 81-92. <<http://www2.aasa.ac.jp/people/smiyata/MiyataMLU2000.pdf>>.
- Miyata, S. (2003). *Wakachi2002 v.2.1*. <<http://childes.psy.cmu.edu/morgrams/Wakachi2002.zip>>.
- Miyata, S. (2004). *Tai-Corpus*. Pittsburgh, PA: TalkBank. 1-59642-057-X.
- 宮田 Susanne, 村木恭子, 森川尋美 (編) (2004). 『今日から使える発話データベース CHILDES 入門』 ひつじ書房.
- Miyata, S. & Naka, N. (2011). *JMOR05 v.1* <<http://childes.psy.cmu.edu/morgrams/jpn.zip>>.
- Miyata, S., Otomo, K., & Nisisawa, H.Y. (2007). Exploring the Developmental Sentence Score for Japanese (DSSJ): A comparison of DSSJ and MLU on the basis of typically developing and delayed children. *The Japanese Journal of Communication Disorders*. 24, 2, 79-88.
- 小椋たみ子・中則 夫・山下由紀恵・村瀬俊樹・マヒューあき (1997) 「日本語獲得児と文法の発達 - CLAN プログラムによる分析」 『神戸大学発達科学部研究紀要』 4, 2, 31-57.
- Park, T.-Z. (1981) The development of syntax in the child with special reference to German. *Innsbrucker Beitrage zur Kulturwissenschaft*. Sonderheft 45.
- Rondal, J. A., Ghiotto, M., Bredart, S., & Bachelet, J.-F. (1987). Age-relation, reliability and grammatical validity of measures of utterance length." *Journal of Child Language* 14. 433-446.
- Rondal, J. A. & DeFays, D. (1978). Reliability of mean length of utterance as a function of sample size in early language development. *The Journal of Genetic Psychology*, 133, 305-306.
- 綿巻徹 (1993) 『日本語 MLU 計算のための発話分割ガイド第2版』 名古屋発達障害研究所.
- 綿巻徹 (1994) 「発話解析プログラム JUPITA による発話の長さとの助詞助動詞獲得研究」 『日本発達心理学会第5回大会発表論文集』 207.
- 綿巻徹 (1999) 『ダウン症児の言語発達における共通性と個人差』 風間書房.
- 綿巻徹・小椋たみ子 (2004) 『日本語マッカーサー乳幼児言語発達質問紙「語と文法」手引』 京都国際社会福祉センター.

補足資料 MLU 計算法の例

1. Tai の 1 才半のサンプル

TAI (対象児)、TMO (母親) と SUU (観察者) の自由会話 (30 分セッションの最初の 100 発話) である。TAI の年齢は 1 才 6 ヶ月 4 日 (t931014.cha)、100 発話にもとづく MLU_w は 1.060、MLU_m は 1.280 である。以下のサンプルでは形態素は「/」、単語は「◁」によってマークされている。左側の数字は①発話の番号 (除かれる発話は 0)、②単語数、③形態素数を表している。

@Begin

				TMO:	どれ?
				SUU:	xxx。
				TMO:	戻したの?
1	1	2	TAI:	切る/◁。	
				SUU:	ん。
0			TAI:	切る。(繰返し)	
				TMO:	un。
0			TAI:	切る。	
				SUU:	これ切るの?
				SUU:	切ってくれる?
				SUU:	パンかなあ?
2	1	1	TAI:	パン/◁。	
				SUU:	ちょうだい!
				SUU:	じゃパンちょうだい!
				SUU:	あ、ありがとう。
				SUU:	全部ばらばらになっちゃったね。
0			TAI:	パン。(繰返し)	
				SUU:	パンを。
				TMO:	ふうん。
3	2	3	TAI:	パン/◁切っ/た/◁。	
				SUU:	これ。
				SUU:	チーズパン。
				SUU:	できた。
				TMO:	チーズパン。
0			TAI:	ボボボボボボ [?]。(不明瞭)	
				TMO:	すごい。
				SUU:	だめ、だめ?
				SUU:	おかしいね。
4	1	1	TAI:	いや/◁。	
				TMO:	いや [?]。
0			TAI:	いや。(繰返し)	
				SUU:	あ、はい。
				SUU:	いやなのね。

(以下省略)

2. Tai の 3 才のサンプル

TAI の年齢は 3 才 0 ヶ月 24 日 (t950504.cha) のセッションである。100 発話にもとづく MLU_w は 2.420、MLU_m は 4.980 である。

@Begin

			TMO:	よしできたぞ。
1	1	3	TAI:	遠/い/でしょう/△?
2	1	2	TAI:	ここ/さあ/△。
			TMO:	うん。
			TMO:	うん。
3	2	5	TAI:	ここ/に/△つなご/ろう/ [=繋げよう] か/△?
			TMO:	うん、そこそこ繋げるよ。
0			TAI:	じゃ、ここ[?]。(不明瞭な発話)
			TMO:	うん、繋げて!
4	3	8	TAI:	でも/△いっぱい/△あ/る/ん/だ/もん/なあ/△。(ため息)
			TMO:	う〜ん。
0			TAI:	いっぱいある xxx。(不明瞭)
0			TAI:	xxx。
			TMO:	いっぱいあるね。
			TMO:	ほら。
0			TAI:	ツクル [?] もんでいや [?]。(不明瞭)
			TMO:	これ?
0			TAI:	ん? (相づち、間投詞、否定、肯定の用言)
			TMO:	[笑]。
0			TAI:	あれ?? (相づち、間投詞、否定、肯定の用言)
			TMO:	あれ??
			TMO:	それここにしようか、こっち。
5	1	1	TAI:	こっち/△?
			TMO:	うん、こっち。
6	1	1	TAI:	こっち/△。
0			TAI:	xxx。
			TMO:	うん、ちょっと悪いかなあ?
			TMO:	ここは?
7	2	5	TAI:	ここ/で/△い/い/じゃない/△?
			TMO:	どれ?
8	2	5	TAI:	ここ/△通行/止め/だ/から/△。
			TMO:	そこ?
9	1	3	TAI:	い/い/よ/△。
10	1	2	TAI:	そこ/は/△?
			TMO:	ここ。
			TMO:	こう[?]?
11	1	1	TAI:	これ/△?
12	1	2	TAI:	違/う/△! (以下省略)