

言語能力の客観的指標の妥当性

——日本人 EFL 大学生の場合——

平 野 絹 枝*

(平成元年10月31日受理)

要 旨

第1言語習得・発達の研究で統語知識・言語能力を測定する指標として、多くの研究者から用いられている T-unit 分析は第2言語習得・学習の研究分野にも広く活用され、T-unit は第2言語習得者の書く能力および言語発達の、妥当性、信頼性の高い指標とされている。本稿では日本人 EFL 大学生の英作文を分析して、日本人 EFL 大学生の英語学力と T-unit 等の客観的指標の関係を明らかにする。

KEY WORDS

EFL index of development	EFL 言語発達の指標	objective measures	客観的指標
error-free T-unit	誤りのない T-unit		
mean T-unit length	T-unit の平均の長さ		
overall language proficiency	総合的言語能力		

1. 第2言語発達の指標

L1 (第1言語, a first language) の言語発達の信頼できる指標として、T-unit を最初に考案したのは、L1 習得研究者の Hunt (1965) である。Hunt (1965, 1970) によれば、T-unit は各文法的単位の最初を大文字に、最後をピリオドか疑問符にした場合に文法的に容認できる最小終結可能単位 (minimal terminable unit) のことである。従って T-unit は1つの主節を含み、その主節に付加される従属節、および、主節の中に埋め込まれたり、主節に付加される非節構造をも含んだ言語単位と定義される (Hunt 1970: 4)¹⁾。その後他の多くの研究者 (O'Donnell, Griffin and Norris 1967, Loban 1976) も T-unit の平均の長さ (T-unit 内の平均語数) が、L1 での文字言語と音声言語の両方の統語的複雑度を示す、客観性と信頼性の高い尺度であることを明らかにした。T-unit の平均の長さは L1 での統語的成熟度の発達、即ち等位節化、次に従属節化、最終的には文の埋め込みや節縮小化を行なう能力によってより多くの情報をより少ない文法的単位の中に取り入れることができる力を反映していると言える。このことは言語能力が高ければ高いほど、長い T-unit を書くことを意味する。

T-unit の平均の長さは英語母国語話者の言語発達の安定した指標として、広く用いられてい

* 言語系教育講座

るが第2言語としての英語 (English as a Second Language, 以下 ESL) の分野でも ESL 学習者の書く能力, および総合的英語学力 (overall language proficiency) を測る指標として利用されるようになった。しかし L2 (第2言語, a second language) 学習者は L1 学習者と違って語彙, 統語, 形態素上の誤りを頻繁におかすという事実から, L2 の言語発達の指標 (L2 index of development) はこの発達上の誤りをある程度考慮したものであるべきという主張が 1970 年代に L2 言語研究者 (Scott and Tucker 1973, Larsen-Freeman and Strom 1977, Larsen-Freeman 1978, Vann 1979) からなされるようになった。その結果, T-unit の平均の長さの他に T-unit を修正した指標, 即ち誤りのない T-unit (error-free T-unit, 以下 EFT) が L2 の作文能力, および総合的言語能力のより妥当な指標であると考えられるようになった。今日では L2 研究では T-unit と EFT の両方が, 統語的成熟度を測る簡便で客観的な指標とみなされている。

第2言語発達の妥当な指標を追及している L2 研究者に Larsen-Freeman がいる。Larsen-Freeman (1978, 1983) は幾つかの広範囲な実証研究を通して, 第2言語発達の指標があれば一つの発達上のものさしとして ESL の総合的言語能力 (overall second language proficiency) を簡便に, 客観的に測定でき, 教師, 教育関係者, 研究者に大きな利益をあたえらうと主張する。Larsen-Freeman (1978, 1983) は, Larsen-Freeman and Strom (1977) の pilot study に引き続いて UCLA の ESL 学習者の作文を分析したところ, 英語学力を最もよく弁別した客観的指標 (objective measures) は, T-unit の平均の長さの他に全 T-unit の中に占める EFT の割合と, EFT の平均の長さであった。つまり学習者の英語力が高くなるにつれて, EFT の平均の長さが長く, かつ全 T-unit に占める EFT の割合も高くなるということを報告している。さらに EFT の平均の長さが, 第2言語発達を測る最も妥当な指標になりうると結論しているのは興味深い。

一方外国語としての英語 (English as a Foreign Language, 以下 EFL) の分野で, 日本人学生の話す能力, 書く能力の客観的指標を扱った実証研究に林 (1979) と門田 (1989) がある。林 (1979) は日本の高校生 10 人を対象に T-unit 分析を行ない, 話す力と書く力の比較を試みている。門田 (1989) は大学生 70 人 (英語専攻者) に対して課題作文を実施し, 作文評価における客観的指標の妥当性を検討している。この2つの研究は日本人学生の英語の口頭能力, 作文能力評価の妥当な指標として EFT をあげており, T-unit の正確さ (EFT の割合) が日本人 EFL 学習者の場合重要な意味を持つことを明らかにしている。

T-unit 分析の利点はその簡便性と客観性にあると言われているが, 一文批判もある。Gaies (1980) は, T-unit 分析は適切さ, 文体上の効果を考慮しないで統語に焦点をおいて分析していること, 伝達能力を必ずしも測定していないこと, 誤りの基準について分析者の間で意見の違いがあることを指摘している。

客観的指標に関する先行研究の結果は一致がみられないが, まとめると ESL/EFL の作文能力, 英語学力の妥当な指標は, 1) 統語的複雑度を示す指標, a) T-unit の平均の長さ (T-unit 1 個あたりの平均語数), b) error-free T-unit (EFT) の平均の長さ (error-free T-unit 1 個あたりの平均語数) と, 2) 正確度を示す指標, 即ち全 T-unit 内の error-free T-unit (EFT) の占める割合 (%) であると言えよう。

2. EFLの英語学力の客観的指標に関する実験

2.1 目 的

ESLでの英語学力測定のためのT-unit分析に関する先行研究に比べ、EFL学習者を対象としたEFLの言語発達の指標を論じた研究(Khered 1983)は少ない。また日本人大学生の弁別力を検討する場合、学力を4群以上に分類して考察した報告は、筆者の知る限りない。従って本研究ではT-unit等の客観的指標が日本人EFL大学生の英語力を測定できる、妥当性のある基準になりうるかどうかを検討する。具体的には英語学力群を3群と4群に分類して次の2点を検証する。1) T-unit等の客観的指標と英語学力の相関はどの位か、2) このような客観的指標が、日本人EFL大学生の英語学力をうまく弁別する基準になりうるかどうか、を明らかにする。

2.2 実 験 方 法

2.2.1 被 験 者

1988, 1989年度の国立大学教養英語受講非英語専攻者(1年生)56名、公立短期大学・国立大学英語専攻者44名(2年生、一部大学院生を含む)計100名。日本人EFL大学生を対象にするため海外在住経験約1年以上の学生4名は分析の対象から除外した。被験者全員、本研究と同じトピックの作文は以前書いたことがないことを、アンケートによって確認した。

2.2.2 材 料

英語学力を測定するため英語学力標準テストとしてのCELT²⁾のForm A(文法テスト、語彙テスト、リスニングテストの3つから構成されている)を利用した(満点300点。各テスト100点満点)。客観的指標の分析として英作文を課した。

2.2.3 手 順

まずCELTを実施して英語学力を測定した。CELTの文法テスト、語彙テスト、リスニングテストの所要時間は各々45分、35分、35分であった。次に制限時間30分の英作文テストを課した。ESL学習者と比較するため、Larsen-Freeman(1978)と同じトピック(“Do you prefer to live in a large city or in a small town?”)を課した。ただし本研究では、被験者に、最初に“I prefer to live in a large city / small town.”から書き出し次にその理由を明確にあげて、主題を展開するように指示した。Perkins(1980)と同様、被験者にはできるだけ多くの英文を書くことを指示した。辞書の使用は禁止した。

2.2.4 分 析 方 法

すべての英作文を、二人の英語母国語話者英語教員と筆者の3人がチェックして誤りの判定を行ない、疑問が残る場合はさらに他の英語母国語話者にもコメントをお願いした³⁾。EFTの分析はLarsen-Freeman(1978, 1983)のあいまい性のない方法を採用し、句読点(punctuation)、つづり(spelling)を含めた、すべての面で誤りのないことをEFTの条件にした。また、文として不完全なもの(fragment, maze)はT-unitの数の計算から除外した(例、One is that ..., Once happen an accident, there become lowdey.)。単語数の計算は次のようにした。(1) 辞書で1語で書かれている単語を、2語として書いた場合、1語と計算(例、every where, sun shine)、(2) 固有名詞(例、New York)、数字(例、100)は1語と計算、(3) 縮小形(例、I'm, won't)は2語と計算、(3) 離して書くべきところハイフンで結んだ場合(例、human-life, every

-day) は2語と計算した。

使用した客観的指標は、(1) 英作文の総語数、(2) T-unit の総数、(3) T-unit の平均の長さ(平均語数) (mean T-unit length=mean number of words per T-unit)、(4) 節の平均の長さ(平均語数)、(5) T-unit 内の節の平均数(clause/T-unit)、(6) 従属節の割合(%)(全ての節の中で従属節の占める割合)、(7) EFT 内の総語数 (total number of words in EFT)、(8) EFT の総数 (total number of EFT)、(9) EFT の平均の長さ(平均語数) (mean length of EFT=mean number of words per EFT)、(10) 全 T-unit 内の EFT の割合 (%) (percentage of EFT)。

2.3 結 果

2.3.1 CELT の平均値と標準偏差

被験者全員の3つのテストの各平均値(満点100点)は、55.25点(最高点76.33, 最低点35.00)、標準偏差は9.30であった。CELTの文法テスト、語彙テスト、リスニングテスト、の信頼係数(KR-21)は各々.74, .75, .64であった。表1は被験者の各学力群のCELTの平均値と標準偏差を示す。表1に示したように各学力レベルのCELTの得点の範囲を設け、被験者をCELT得点に基づいてレベル1(下位群)、レベル2(中位群)、レベル3(上位群)と3群に分類した。さらに客観的指標の精度をみるために、同じ被験者を4群に分類して考察した。

表1 3群と4群におけるCELTの平均得点と標準偏差

3 groups					4 groups				
level	n	得点範囲	\bar{X}	SD	level	n	得点範囲	\bar{X}	SD
1	20	35.0—48.9	41.67	3.79	1	17	35.0—45.9	40.63	3.11
2	62	49.0—62.9	55.52	3.65	2	40	46.0—56.9	52.63	2.75
3	18	63.0—76.9	69.41	3.64	3	31	57.0—67.9	60.43	2.91
					4	12	68.0—78.9	71.31	2.76

2.3.2 CELT と客観的指標の相関係数

表2が示すように本研究で検出された、CELTと客観的指標の相関は他のテスト(TOEFL, プレイスメントテスト等)を外部基準テストとして用いた先行研究の結果と同様、決して高くないが、T-unitの総数と節の平均の長さの2つの指標を除いてすべてCELTの得点と1%以下の危険率で有意な相関があった。5台と4台の有意な相関係数(ピアソンの積率相関係数)を示したのは、T-unitの平均の長さ、EFT内の総語数、EFTの総数、全T-unit内のEFTの割合、であった。Larsen-Freeman(1978, 1983)が最も妥当性の高い指標として提唱している、EFTの平均の長さは有意であるが、.362という低い相関係数であった。Hunt(1965, 1970)他が有効な指標と提唱している節の平均の長さはCELTと有意な相関が認められず、T-unit内の節の平均数、全節内の従属節の割合はCELTと低い相関を示した。

2.3.3 客観的指標の弁別力

さらに客観的指標が3群と4群の各学力レベルをうまく弁別するかどうか調べるために分散分析を行なった結果をまとめたのが表3と表4である。分散分析の結果、3群と4群間に有意差が検出された場合、さらにシェッフエの法によって各学力レベル間の有意差を調べた。表3と

表2 CELT と客観的指標の相関

英作文の総語数	.316**
T-unit の総数	.025
T-unit の平均の長さ (平均語数)	.487***
節の平均の長さ (平均語数)	.092
T-unit 内の節の平均数	.278**
従属節の割合	.327***
EFT 内の総語数	.519***
EFT の総数	.458***
EFT の平均の長さ (平均語数)	.362***
全 T-unit 内の EFT の割合	.464***

**p<.01

***p<.001

表4 から次のことが明らかになった。

3つの学力群の場合

- (1) T-unit の総数と、節の平均の長さの2つの指標を除いた8つのすべての指標では3つの学力群間に有意差があることが分った。従ってこれら8つの指標がESLの学力を弁別するのに有用な統語的指標として働くと言える。
- (2) すべての隣接レベル (level 1 と level 2, level 2 と level 3) 間を有意に弁別できた指標はT-unit の平均の長さ、EFT 内の総語数であった。
- (3) 英作文の総語数、T-unit 内の節の平均数、EFT の総数、の3つの指標は学力の隣接レベル間をすべて有意に弁別しなかった。

4つの学力群の場合

- (1) T-unit の総数、節の平均の長さ、T-unit 内の節の平均数を除いた7つの指標の学力間に有意差があった。
- (2) すべての隣接レベル (level 1 と 2, 2 と 3, 3 と 4) 間を有意に弁別できた指標は一つもなかった。
- (3) T-unit の平均の長さ、EFT 内の総語数、全 T-unit 内の EFT の割合は、他の指標より多くの学力群間 (6対のうち、4,5対) を有意に弁別した。
- (4) 英作文の総語数は、最上位 (level 4) と最下位 (level 1) の学力間しか有意に弁別せず、また EFT の平均の長さは、多くの学力間を有意に弁別しなかった。

以上3群と4群に分類した結果をまとめると、

- (1) 4つの学力群に分類するとすべての隣接レベル間を有意に弁別する指標は一つもなくなり、3つの学力群の場合より学力を弁別しにくくなる。このことから、各英語力のレベルに幅がある方が英語力を弁別しやすくなると言える。
- (2) 10種の客観的指標の中で、T-unit の平均の長さ、EFT 内の総語数、全 T-unit 内の EFT の割合が、すべての学力群間、または他の指標に比べてより多くの学力群間を有意差をもって弁別する。

表3 3つの学力群における客観的指標の分散分析の結果

指 標	level	\bar{X}	SD	F 値	post hoc analysis	
英作文の総語数	1	139.25	41.00	3.768	1 vs. 2	NS
	2	160.37	44.67	p<.05	1 vs. 3	p<.05
	3	180.06	50.91		2 vs. 3	NS
	全体	159.69	46.89			
T-unit の総数	1	15.00	4.81	0.160		
	2	15.29	4.28	NS		
	3	14.61	4.99			
	全体	15.11	4.53			
T-unit の平均の長さ	1	9.23	1.58	18.234	1 vs. 2	p<.05
	2	10.52	1.80	p<.001	1 vs. 3	p<.001
	3	12.59	1.47		2 vs. 3	p<.001
	全体	10.64	2.00			
節の平均の長さ	1	7.58	1.42	1.026		
	2	8.16	3.40	NS		
	3	8.90	1.23			
	全体	8.18	2.83			
T-unit 内の節の平均数	1	1.23	0.19	3.790	1 vs. 2	NS
	2	1.35	0.24	P<0.5	1 vs. 3	p<.05
	3	1.43	0.21		2 vs. 3	NS
	全体	1.34	0.23			
従属節の割合(%)	1	17.04	12.05	6.380	1 vs. 2	p<.05
	2	25.26	10.38	p<.01	1 vs. 3	p<.01
	3	28.80	9.52		2 vs. 3	NS
	全体	24.25	11.26			
EFT 内の総語数	1	24.15	15.64	10.345	1 vs. 2	p<.05
	2	41.42	25.47	p<.001	1 vs. 3	p<.001
	3	62.89	34.69		2 vs. 3	p<.05
	全体	41.83	28.47			

指 標	level	\bar{X}	SD	F 値	post hoc analysis
EFT の総数	1	3.10	1.84	6.316	1 vs. 2 NS
	2	4.68	2.47	p<.005	1 vs. 3 p<.005
	3	6.00	3.11		2 vs. 3 NS
	全体	4.60	2.65		
EFT の平均の長さ	1	7.83	1.77	8.525	1 vs. 2 NS
	2	8.76	2.01	p<.001	1 vs. 3 p<.001
	3	10.40	1.73		2 vs. 3 p<.01
	全体	8.87	2.08		
全 T-unit 内の EFT の割合(%)	1	21.81	11.91	7.189	1 vs. 2 p<.05
	2	31.18	15.62	p<.005	1 vs. 3 p<.005
	3	39.70	12.06		2 vs. 3 NS
	全体	30.84	15.37		

注：NS=non-significant (有意差のないことを表す)

表4 4つの学力群における客観的指標の分散分析の結果

指 標	level	\bar{X}	SD	F 値	post hoc analysis
英作文の総語数	1	137.24	43.60	3.803	1 vs. 2 NS
	2	156.90	38.14	p<0.5	1 vs. 3 NS
	3	162.26	48.02		1 vs. 4 p<.05
	4	194.17	53.63		2 vs. 3 NS
	全体	159.69	46.89		2 vs. 4 NS 3 vs. 4 NS
T-unit の総数	1	15.00	5.08	0.095	
	2	15.10	4.10	NS	
	3	14.94	4.37		
	4	15.75	5.34		
	全体	15.11	4.53		

指 標	level	\bar{X}	SD	F 値	post hoc analysis	
T-unit の平均の長さ	1	9.10	1.53	9.488	1 vs. 2	NS
	2	10.49	1.81	p<.001	1 vs. 3	p<.05
	3	10.89	1.84		1 vs. 4	p<.001
	4	12.65	1.67		2 vs. 3	NS
	全体	10.64	2.00		2 vs. 4	p<.01
					3 vs. 4	p<.05
節の平均の長さ	1	7.55	1.49	0.766		
	2	8.37	4.15	NS		
	3	7.93	1.05			
	4	9.03	1.30			
	全体	8.18	2.83			
T-unit 内の節の平均数	1	1.22	0.19	2.295		
	2	1.34	0.24	NS		
	3	1.38	0.22			
	4	1.42	0.23			
	全体	1.34	0.23			
従属節の割合(%)	1	16.21	12.56	3.947	1 vs. 2	p<.05
	2	25.21	9.48	p<.05	1 vs. 3	p<.05
	3	26.01	11.02		1 vs. 4	p<.05
	4	27.91	10.23		2 vs. 3	NS
	全体	24.25	11.26		2 vs. 4	NS
					3 vs. 4	NS
EFT 内の総語数	1	23.00	15.04	14.127	1 vs. 2	NS
	2	33.13	18.89	p<.001	1 vs. 3	p<.005
	3	50.42	26.91		1 vs. 4	p<.001
	4	75.33	36.14		2 vs. 3	p<.05
	全体	41.83	28.47		2 vs. 4	p<.001
					3 vs. 4	p<.05
EFT の総数	1	3.00	1.91	9.561	1 vs. 2	NS
	2	3.88	2.00	p<.001	1 vs. 3	p<.05
	3	5.45	2.49		1 vs. 4	p<.001
	4	7.08	3.23		2 vs. 3	NS
	全体	4.60	2.65		2 vs. 4	p<.005
					3 vs. 4	NS

指 標	level	\bar{X}	SD	F 値	post hoc analysis
EFT の平均の長さ	1	7.79	1.64	4.814	1 vs. 2 NS
	2	8.66	2.10	p<.005	1 vs. 3 NS
	3	9.10	1.85		1 vs. 4 p<.01
	4	10.53	1.99		2 vs. 3 NS
	全体	8.87	2.08		2 vs. 4 p<.05
			3 vs. 4 NS		
全 T-unit 内の EFT の割合(%)	1	20.91	11.27	10.275	1 vs. 2 NS
	2	26.24	13.49	p<.001	1 vs. 3 p<.005
	3	37.33	14.84		1 vs. 4 p<.001
	4	43.47	11.64		2 vs. 3 p<.05
	全体	30.84	15.37		2 vs. 4 p<.005
			3 vs. 4 NS		

注：NS=non-significant (有意差のないことを表す)

2.4 考 察

以上の結果をまとめると次の事が言える。

- (1) 本研究では、CELT 得点との相関係数と、3, 4 段階の各学力レベルの弁別の度合いから日本人 EFL 大学生の英語力を測る、より有力な指標は、① 統語的複雑度を示す「T-unit の平均の長さ」、② 「EFT 内の総語数」、③ 正確度を示す「全 T-unit 内に占める EFT の割合」であると、結論できる。つまり、英語学力が高くなるにつれて、T-unit の平均の長さがより長く、EFT 内の単語数がより多く、EFT の割合もより高くなる傾向があると言えよう。Khered (1983), Perkins (1980) の研究結果でも、EFT 内の総語数が最も有効な指標のひとつであることを明らかにしており、計算の簡便性からしてこれが客観性のある、容易に利用できる指標の一つであることに注目すべきであろう。
- (2) EFT に関する指標は、CELT のテストとすべて有意な相関を示し 3, 4 群の学力間に有意差が認められた。従って、EFT に関する指標は、EFL 学習者の英語学力測定の有効な指標と言えよう。
- (3) L2 学習者の言語能力の最も有効な指標とされている EFT の平均の長さは、他の EFT に関する指標より、CELT 得点との相関は低く、4 つの学力群の場合、6 対のうち 2 対しか有意に弁別しなかった。このように EFT の平均の長さは日本人 EFL 大学生の場合必ずしも妥当性の最も高い指標にはならない。
以上の 2) と 3) は林 (1979), Khered (1983), 門田 (1990) の結果からも、ほぼ言えることである。
- (4) L1 学習者の有効な指標と提唱されている節の平均の長さ、T-unit 内の節の平均数は、CELT の得点と有意な相関がないか、低い、さらに 4 群の学力間に有意差をもって弁別しな

い。このことから「T-unit の平均の長さ」を除いて、誤りを無視した指標は日本人 EFL 大学生の英語学力を示す妥当な指標とは必ずしも言えない。

- (5) 本研究では、T-unit の平均の長さも、日本人 EFL 大学生の英語学力の妥当な指標になりうるということが明らかになった。この T-unit の平均の長さが、ESL/EFL での作文能力、英語学力の妥当な指標になり得るかどうかについては先行研究の結果に一致がみられない。門田 (1990) では T-unit の平均の長さは日本人大学生の有効な指標とはならなかった。このように結果が異なるのは、被験者集団の学力の幅、T-unit の測定方法、作文テストの実施方法、作文のトピック、ジャンル、外部基準テストの種類の違い、L1 等の要因が関係しているのかもしれない。
- (6) CELT の得点と客観的指標の相関は、他の英語標準テスト (TOEFL, ミシガンテスト) を用いた先行研究と同様、高いとは必ずしも言えない。この種の外部テストは概して言語操作能力 (linguistic competence) を測定する discrete-point test で、作文は総合的言語運用能力を要求するためであろう。

3. おわりに

ESL/EFL 分野での T-unit 等の客観的指標の限界は、隣接レベルの学力間の弁別が期待したほど明確でないという点である。また、英語学力の複雑な定義の問題もさることながら、T-unit 分析には、T-unit, fragment の測定方法、error の判定、被験者集団の特質、伝達能力への配慮が欠けていることなど、いくつかの問題が残り、改善すべき点がある。さらに英語力以外の他の要因 (作文のトピック、ジャンル、難易度、英語経験等) との関係も考慮し、今後さらに検討を重ねていく必要がある。

ESL/EFL の言語発達の妥当な指標は、学習者の言語能力を量的に測る際の、客観的なひとつの補助的指標 (めやす) として用いることの方がよく、あまり過大に期待するのは慎むべきであろう。

注

- 1) 重文 (例, I have an uncle and his name is Roger.) は T-unit 2 つ, 複文 (例, I have an uncle, whose name is Roger.) は T-unit 1 つである。
- 2) Harris, D.P. & L.A. Palmer. 1986. A Comprehensive English Language Test for Learners of English (McGraw-Hill).
- 3) Ms. Leslie Robertson, Mr. Fred Durbin, Ms. Lynn Mucciacciaro には厚く謝意を表す次第である。

参 考 文 献

- Arthur, B. (1979). Short-term changes in EFL composition skills. In C. Yorio et al. (Eds.), *On TESOL '79* (pp. 330-342). Washington, D.C.: TESOL.
- Cooper, T. (1976). Measuring written syntactic patterns of second language learners of German. *Journal of Educational Research*, 69 (5), 176-183.
- Crowhurst, M. (1980). Syntactic complexity and teacher's quality ratings of narrations and arguments. *Research in the Teaching of English*, 14, 223-231.
- Dvorak, T. (1987). Is written FL like oral FL? In B. VanPatten et al. (Eds.), *Foreign language learning* (pp. 79-91). NY: Newbury House.
- Flahive, D. & Snow, B. (1980). Measures of syntactic complexity in evaluating ESL compositions. In J. Oller & K. Perkins (Eds.), *Research in language testing* (pp. 171-176). Rowley, MA: Newbury House.
- Gaies, S. (1980). T-unit analysis in second-language research: applications, problems, and limitations. *TESOL Quarterly*, 14 (1), 53-60.
- 林 日出夫 (1979). 「外国語習得における話す力と書く力のバランスについて」『語法研究と英語教育』山口書店, 第2号, 52-65.
- Hirano, Kinue (1989). Research on T-unit Measures in ESL. *Bulletin of Joetsu University of Education*, 8, sect. 2, 67-77.
- Homburg, T. (1984). Holistic evaluation of ESL compositions: can it be validated objectively? *TESOL Quarterly*, 18 (1), 87-107.
- Ho-Peng, L. (1983). Using T-unit measures to assess writing proficiency of ESL students. *RELC Journal*, 14 (2), 35-43.
- Hunt, K. (1965). *Grammatical structures written at three grade levels*. NCTE Research Report, No. 3. Urbana, IL: The National Council of Teachers of English.
- Hunt, K. (1970). Syntactic maturity in schoolchildren and adults. *Monographs of the Society for Research in Child Development*, 35, serial No. 134, 1-67.
- Kameen, P. (1979). Syntactic skill and ESL writing quality. In C. Yorio et al. (Eds.), *On TESOL '79* (pp. 343-350). Washington, D.C.: TESOL.
- 門田修平 (1990). 「作文評価における客観的指標の妥当性」『英語・英文学新潮』ニューカレント・インターナショナル, 339-350.
- Khered, M.O.H. (1983). *Measuring the syntactic maturity in the written English of Arab studentns at four proficiency levels and establishing an EFL index of development*. Unpublished Ph. D. dissertation, Univ. of Kansas.
- Larsen-Freeman, D. (1978). An ESL index of development. *TESOL Quarterly*, 12 (4), 439-447.
- Larsen-Freeman, D. (1983). Assessing global second language proficiency. In H.W. Seliger & M.Long (Eds.), *Classroom-oriented research in second language acquisition* (pp. 287-304). Rowley, MA: Newbury House.
- Larsen-Freeman, D. & Strom, V. (1977). The construction of a second language acquisition

- index of development. *LL*, 27 (1), 123-134.
- Loban, W.D. (1976). *Language development : kindergarten through grade twelve*. NCTE Research Report 18. Urbana, IL : National Council of Teachers of English.
- Monroe, J.H. (1975). Measuring and enhancing syntactic fluency in French. *The French Review*, XLVIII (6), 1023-1031.
- Ney, J.W. (1966). Review of grammatical structures written at three grade levels (Hunt K. W.). *LL*, 16 (3-4), 230-235.
- O'Donnell, R.C., Griffin, W.J. & Norris, R.S. (1967). *Syntax of kindergarten & elementary school children : A transformational analysis*. NCTE Research Report 8. Urbana IL : National Council of Teachers of English.
- O'Hare, F. (1973). *Sentence combining : improving student writing without formal grammar instruction*. NCTE Research Report 15. Urbana, IL : National Council of Teachers of English.
- Perkins, K. (1980). Using objective methods of attained writing proficiency to discriminate among holistic evaluations. *TESOL Quarterly*, 14 (1), 61-69.
- Perkins, K. (1983). On the use of compositions scoring techniques, objective measures, and objective tests to evaluate ESL writing ability. *TESOL Quarterly*, 17 (4), 651-671.
- Potter, R. (1967). Sentence structure and prose quality. *Research in the Teaching of English*, 1 (1), 17-28.
- Scott, M. & Tucker, G.R. (1974). Error analysis and English-language strategies of Arab students. *LL*, 24 (1), 69-97.
- Witte, S. (1983). The reliability of mean T-unit length : some questions for research in written composition. In A. Freedman, I. Pringe & J. Yalden (Eds.), *Learning to write : first language/second language* (pp. 171-177). NY : Longman.
- Vann, R. (1979). Oral and written syntactic relationships in second language learning. In C. Yorio et al. (Eds.), *On TESOL '79* (pp. 322-329). Washington, D.C. : TESOL.

The validity of Objective Measures of Japanese EFL Students' Proficiency

Kinue HIRANO

ABSTRACT

The T-unit has been successfully used in both L1 and L2 research. This study reports on a piece of research which examines the validity of ten objective measures of language development in EFL. Mean T-unit length, the number of words in error-free T-units, and the percentage of error-free T-units discriminated very well among Japanese EFL students at different levels of proficiency.