TYPE AND QUANTITY OF FEEDBACK USE AS A MEASURE OF LANGUAGE PROFICIENCY

Jens Allwood

ML U as a measure of language proficiency

A rough solution to the problem of assessing developing proficiency is to consider the development of the so-called mean length of utterance (MLU) for a certain learner, cycle by cycle. This measure, although it has many problems (see below) has often been used to give a very rough idea both of differences in proficiency and of the extent to which acquisition is taking place. In this chapter, an utterance is defined as a stretch of uninterrupted speech produced by one speaker bounded by silence or by contributions from other speakers.

The performances of twenty learners, whose names are given below, are examined in this study. The activities studied are conversations (A), role plays (J) and play scenes (C), which are described in chapter 6. Further discussion of the internal dynamics of these activities can be found in Volume 11:11.2. These performances are compared with those of a small group of native speakers (the 'controls') in identical activities. This group is composed of two working-class native speakers of English (Martin and Sheila), and of Swedish (Adam and Eva). In addition, Marl and Nora performed these activities in their native language as well as in Swedish.

In Table 8.1 we reproduce the results concerning mean length of utterances for learners grouped according to pairs of source and target languages. The table contains the means for the individuals, as calculated over two activity occurrences in cycle 1 and cycle 3, and the differences between the means in cycle 3 and cycle 1 in order to give an idea of the change that has taken place between the two first and two last recordings.[1] The table reports on individual learners rather than on SL-TL aggregations since there are too few individuals and too many potential differences between them really to justify anything but a report on the individual level. Table 8.1 shows that there is indeed considerable individual variation both between learners of a given target language and between learners sharing a source language. Compare Berta and Alfonso who are both Latin American Spanish-speaking learners of French, or Lavinia and Andrea who are both Italian-speaking learners of English. It is not unlikely that these differences reflect a difference in target language proficiency at the beginning of the period of observation and that there are different rates of acquisition for different learners. The differences between the learners with regard to the MU' for cycle 1 results in the following ordering, which we provisionally take to be a first, rough indicator of differences in proficiency in cycle 1: Ravinder *(2.0),* Çevdet *(2.4),* Ergün *(2.9),* Fatima *(3),* Madan *(3.2),* Leo *(3.3),* Fernando *(3.5),* Man (3.7), Berta *(3.7),* Zahra *(3.7),* Mahmut *(3.8),* Ilhami *(3.8),* Andrea *(3.9),* Marcello *(4),* Tino *(4.3),* Mohamed *(4.4),* Abdelmalek *(4.4),* Lavinia *(4.9),* Nora *(7),* Alfonso *(15).* A general point to notice is that all the learners except one - Mohamed - show an increase in MLU between cycles 1 and *3.*

Table 8.1 *Development of MLI.); cyclic mean per activity occurrence (Learners grouped by source and target languages; twenty learners)*

			Cl	C3	C1-C3
Sw	Fl	Marl	3.7	6.6	2.9
		Leo	3.3	5.1	1.8
	Sp	Nora	7.0	9.2	2.2
		Fernando	3.5	5.8	2.3
Fr	Sp	Berta	3.7	17.2	13.5
		Alfonso	15.0	17.5	2.5
	Ar	Zahra	3.7	5.5	1.8
		Abdelmalek	4.4	7.9	2.5
Du	Ar	Fatima	3.0	3.4	0.4
		Mohamed	4.4	3.5	-0.9
	Tu	Ergün	2.9	3.7	0.8
		Mahmut	3.8	4.8	1.0
Ge	Tu	Çevdet	2.4	6.2	3.8
		Ilhami	3.8	5.9	2.1
	It	Marcello	4.0	4.3	0.3
		Tino	4.3	7.2	2.9
Eng	it	Lavinia	4.9	9.0	4.1
		Andrea	3.9	5.0	1.1
Pu		Ravinder	2.0	4.0	2.0
		Madan	3.2	3.7	0.5
Total mean:			4.3	6.8	2.5

[1] See chapter 5.3 for the data collection

In sum the data, both on an individual level and on a more aggregated level show an increase in MU) which is compatible with a gradual increase in individual linguistic proficiency. Comparing the learners with the controls we find that the overall mix for the controls is 10.7 as compared with a total overall MLI.) of 5.5 for the learners. This indicates that MLI.) is a measure which is sensitive to whether one is learning a language or not and therefore perhaps can be used as an indicator of differences in language acquisition.

As we have already said, all learners except Mohamed show a higher MLI.) in the third cycle than in the first, which is surprising since Mohamed has been characterised as a fast learner of Dutch. If we accept this assessment of Mohamed as correct, it is not obvious how one should reconcile it with his MI.) decrease in the third cycle. One possibility is that MLU does not reflect linguistic proficiency. Another possibility is that MLU only indirectly reflects linguistic proficiency and that we should try to locate other factors that might also be reflected by MU). In Mohamed's case, we know that he was the one learner whose motivation fell away sharply during the third cycle, and he was minimally co-operative during project activities. The low MLU may reflect this fact. To the extent that MI.) reflects general proficiency, we see that it does so only given an assumption of a constant level of co-operation between researchers and informants.

Some of the expectable differences in MLU scores can perhaps be elucidated by reference to differences between source and target languages. In particular, lexical similarity (and to some extent grammatical similarity) between source and target languages is likely to facilitate fluency in the learners' TL communication and, thus, probably also MLU. For example, Punjabi and English are somewhat less closely related than Swedish and Spanish, and are considerably less closely related than Spanish and French. This might, ceteris paribus, mean that it is easier for Spanish speakers to learn French and Swedish than it is for Punjabi speakers to learn English. This might further mean that one should expect Punjabi learners of English to have lower MLUS than Spanish learners of Swedish, who in turn should have lower MLIJS than Spanish learners of French. Table 8.1 supports these expectations.[2]

However, the relation between utterance length and complex syntactic-semantic cohesion is not one-one. So even though some types of syntactic-semantic cohesion require long utterances, there are also short utterances with great complexity with regard to syntacticsemantic cohesion and long utterances with little complexity with regard to such cohesion. The relation between MLU and proficiency is therefore an approximation and should wherever possible be supplemented by other measures. We turn now to one such measure; feedback use.

Feedback as a supplementary measure of language proficiency

One of the areas under investigation in our study is feedback. Feedback can briefly be characterised as specialised linguistic mechanisms (for our present needs, words) which enable speech partners to give and elicit information on the basic dialogic requirements of contact, perception, understanding and attitudes. The phenomenon ('narrow feedback') is briefly described in chapter 4.2, and discussed in detail in Volume 11:11.2. This area not only has interest, in its own right, as an aspect of linguistic communication but it can perhaps also provide data that can be used as a measure of linguistic proficiency to supplement MLU. The reason for this is that the relative share of narrow feedback in a given learner's output, in general, seems to decrease as the learner increases his linguistic proficiency. Possible explanations for this decrease are that:

(i)   narrow feedback is among the first linguistic means available to a learner;
(ii)   narrow feedback can be used to substitute for or to accomplish other linguistic functions; and
(iii)   as the learner acquires other linguistic means he will use these to accomplish what he earlier accomplished by feedback.

The two main measures we have used to get an idea of the relative share of narrow feedback expressions in the learners' linguistic output are FRU (relative share of feedback-containing utterances in relation to total number of utterances in an activity occurrence) and FBW (relative share of feedback words in relation to total number of words in an activity occurrence). Using these two measures, Table 8.2 gives us an idea of the relative percentage amount of feedback expressions for the different learners. The table contains the cyclic means for the individuals, and the per cent unit difference between the means in cycle 3 and cycle 1. As in Table 8.1, what Table 8.2 presents could be viewed as a number of case studies in tabular form.

---

[2] 1t should be said that AIfonso was an exceptionally talkative and sociable person, who already a fairly rich repertoire at the first recordings analysed here.

Table 8.2 FBU find FBW, mean relative shares per learner and cycle (twenty learners,)

			FBU			FBW		
			Cl	C3	Cl-3	Cl	C3	C1-3
Sw	Fi	Mail	74	56	-18	23	12	-11
		Leo	81	79	- 2	42	26	-16
	Sp	Nora	74	63	-11	21	5	- 6
		Fernando	65	59	- 6	29	15	-14
Fr	Sp	Berta	69	57	-12	31	6	-25
		Alfonso	63	56	- 7	11	7	- 4
	Ar	Zahra	79	63	-16	33	15	-18
		Abdelmalek	69	53	-16	34	17	-17
Du	Ar	Fatima	67	55	-12	27	20	- 7
		Mohamed	63	73	10	19	28	9
	Tu	Ergün	54	66	12	23	27	4
Mahmut		69	80	11	27	27	0	
Ge	Tu	Çevdet	64	53	-11	35	11	-24
		Ilhami	58	51	- 7	25	11	-14
	It	Marcello	57	55	- 5	26	17	- 9
		Tino	51	53	2	22	11	-11
Eng	ft	Lavinia	80	72	- 8	27	17	-10
		Andrea	51	74	23	27	24	- 3
	Pu	Ravinder	61	84	23	40	37	- 3
		Madam	76	75	- 1	43	42	-I
Total mean pr cycle			66	64	-2	28	21	-7
Total mean controls			62			16		

The table shows that there is a small overall FRU decrease and a somewhat greater FDW decrease. The trend is clearer for FBW than for FRU. This judgement is motivated not just by the numerical difference, visible in the table, but also by a consideration of the base for the calculation of the relative shares of FRU (9,772 utterances) and FHW (49,474 words). Although both measures rest on secure grounds, we see that the absolute numbers required for a decrease in the relative share of FRW (as measured in percentages) are much greater than those required for a decrease in the relative share of FRU.

A comparison with the total means of the controls (native speakers' data) for FRU and FBW supports the analysis we have made of the trends for learners concerning FBU and FBW. The controls have both a lower mean FRU score and a lower mean mw score than the majority of the learners exhibit, even in cycle 3. This means that high initial and Successively decreasing scores of FRU and FBW can perhaps be taken as something which is indicative of acquisition. We will return to why this might be so below.

Let us now look a little more carefully at the FRU and FBW scores. We observe that six learners (Mohamed, Ergün, Mahmut, Tino, Andrea and Ravinder) increase their FRU from cycle 1 to cycle 3, while only 2 learners (Mohamed and Ergün) increase their FBW rate. Mohamed's lack of co-operation in cycle 3 is manifested here by his use of minimal feedback.

Table 8.2 does not allow for any statistically sound inferences to be drawn. It can, however, be used to look for trends which can then lend support to certain hypotheses. The data can also be used to check for compatibility with and to gain initial support for hypotheses which can be proposed on partly independent grounds.

The following are possible hypotheses which can be seen as an attempt to specify the three explanations put forth initially in this section:

(i)   FB words often have simple phonological structure. They can therefore be learned early and used fairly easily.
(ii) There is a constant need and use of feedback in most types of spoken interaction. FB words are therefore usually available in the spoken input which the learners are exposed to and they have a high need to make use of this input.
(iii) Initially, basic feedback functions and basic linguistic feedback mechanisms can be used to substitute for other more specific linguistic functions.
(iv) Initially, feedback functions are also used by the learner as a means for language acquisition.
(v)   The reasons given in (iii) and (iv) but not in (i) and (ii) can be expected to diminish in importance as the learners proficiency increases.

The data in Table 8.2 seems compatible with these assumptions. The total FBU rate remains fairly constant with a slight decrease. Both learners and controls have a high FBU rate, with an average difference of only 4 per cent.

This result can be taken as weak support for the hypothesis that there is a constant and fairly high need of feedback for everyone and that this need is slightly higher for language learners. However, the fact that there is a fairly high degree of variability between learners with regard to FBU (for example, six learners increase their rate from cycle 1 to cycle 3) seems to indicate that FRU is sensitive to factors other than proficiency. Such factors could, for example, include motivation and the kind of activity in which the learner is engaged.

The relationship between FBU, FBW, MLU and proficiency
Turning to the FRW rate, we should first note that there is a conceptual relationship between FRU and FBW. The FRU rate is the relative share of utterances containing feedback words and the FBW rate is the relative vocabulary share of the FB words. Thus, the two measures have a common base in the FR words, which in the FBW measure are used directly in relation to the total vocabulary of a learner in a specific activity occurrence, and in the FRU measure are lumped together relative to the utterance containing them, and then compared with the total number of utterances of the learner in a certain activity occurrence. There must, therefore, on analytical grounds, be a relationship between the two measures.

In spite of this relationship, the two measures also show some differences which are interesting to note. The FBW rate shows a more consistent decrease than the FBU rate (only two learners show an increase). This could be interpreted in the following way: whereas the FBU rate reflects a fairly constant need for feedback which can be affected by such things as motivation and type of activity, the FBW rate, although also sensitive to the same factors, shows a more direct relation to language acquisition. This interpretation is supported both by the fact that the difference between controls and learners is slightly larger for FBW than for FBU and by the relatively uniform decrease in FBW for the majority of the learners.

The reason that FBW should show a closer connection to language acquisition than FBU is that the former, according to our hypothesis, should initially be usable both as an instrument for language acquisition and to substitute for linguistic functions other than feedback. As acquisition proceeds the learner will acquire more specific linguistic means for these functions. Since the need for feedback is fairly constant the changes should affect FBW (the relative share of FR words of the total vocabulary) more directly than FBU (the relative share of feedback-containing utterances). The FBU rate can remain high and permit many functions, other than the basic feedback functions, to be performed sequentially or simultaneously with feedback, while this can not be the case for the FBU rate, which will immediately be affected by the addition of other functions when they are lexically encoded.

There is a negative correlation between MLU and FBW (Pearson's product moment: -.67), that is, when MLU is high, FBW tends to be low. The longer utterances a learner has, the smaller the share of FB words in his vocabulary (although it is always possible to give FB by repeating the interlocutor's words; see below). As one learns to produce longer utterances, one also learns to make greater use of non-feedback vocabulary. In this connection, it is interesting to note that the only learners who have a higher MLU score in cycle 3 than the controls, that is, the Spanish-French learners, are also the learners who have the lowest FOW rate. The connection between high MLU, low FBW, and proficiency receives further support from a comparison of learners and controls with regard to the correlation MLU-FBW. The controls exhibit an even stronger negative correlation than the learners (Pearson's product moment: -.76 versus -.67).

Another clear way of demonstrating the fact that feedback seems to have a special function in language acquisition is to compare linguistic data for one and the same learner, in one and the same activity in SL and TL. The data exhibited in Table 8.3 come from the only two informants for whom we have control data of this type available.

Table 8.3 comparison of MLU, Fill) and FIIW in SL and TL (Mari and Nora)

| | SL | | | TL | | |
	MLU	%FBU	%FBW	MLUI	%FBU	%FBW
Marl	13.7	56.0	7.5	5.1	63.9	20.2
Nora	18.6	547	5.9	8.6	71.2	18.4

The status of MLU as a measure of proficiency is corroborated. For both learners there is a mean difference of almost ten words per utterance between speaking their first language and speaking the target language. The same differences are present if we observe the share of feedback-containing utterances and feedback words. We see that the share increases dramatically, especially in relation to feedback words when learners are using the target language. There is also an increase, although a smaller one, in the amount of feedback-containing utterances in the TL data.

Farther refinements
Over and above the use of FBW (and to some extent FBU) as an indicator of linguistic proficiency, two other aspects of feedback also show a correlation to changes occurring in language acquisition, namely, the functions of feedback words and repetitions as feedback.

Let us therefore take a look at how the two main FB functions - giving and elicitation - are distributed with regard to utterance status and utterance position. The data base for the learners is 7153 tokens used with Fl) function. The developmental patterns in relation to the crossing of the two dimensions is shown in Table 8.4. *Initial,* medal, *final* are utterance positions, while single indicates that an utterance has no other function than feedback. The functions are: (i) pure feedback giving (FBG), (ii) pure feedback elicitation (FBE) and (iii) utterances which have both a feedback giving and a feedback eliciting function (FBG/FBE).

Table 8.4 *Development of FBG, FBE and FBG/FBE in relation to utterance status and utterance position (percentage of FB units, twenty learners)*

	Pure FBG			Pure FBE			FBG/FBE		
	Cl	C3	C1-3	Cl	C3	C1-3	Cl	C3	C1-3
Single	59.5	51.1	-8.4	0.8	0.8		9.0	3.1	-59
Initial	23.4	32.6	9.2	0.1	0.3	0.2	0.9	5	-0.4
Medial	0.8	1.7	0.9	0.1	0.6	0.5		0.1	0.1
Final	1.3	1.7	0.4	0.4	2.9	2.5	0.1	0.2	0.1
% Total FB units	83.0	87.4	2.4	3.6	6.7	3.1	11.4	5.9	-5.5

The decrease of the relative shares of singles - utterances consisting only of FB material - is clearly indicated. We see that it concerns singles as pure FBG and singles used in a vague way (FBG/FBE). The singles that are used for pure elicitation, on the other hand, retain a small but constant share. There is a small increase in all utterance positions, above all in initial position, of pure FBG. There is a concurrent but much smaller increase in all positions, especially final, or pure FBE and, finally there is a decrease among singles and in initial position of FBG/FBE.

The data supports the following generalisations about language acquisition and feedback: (i) that there will be a gradual functional differentiation of giving and eliciting; (ii) that both of these functions will be manifested with increasing complexity in relation to other linguistic material (that is, occur embedded in utterances rather than as single-word utterances); (iii) that 'pure FBG' will tend to be manifested either in single utterance or in initial utterance position, where the initial utterance position will gradually take over some of the 'pure giving' function from the singles; (iv) that 'pure FBE' will tend to be manifested in single utterances or in final utterance position and (v) that the vague category, FBG/FBE, will be manifested primarily as singles and that it will show a decrease from cycle 1 to cycle 3 connected with the learners' increasing differentiation of feedback functions. These generalisations will be further discussed in Volume 11:11.2.

Besides the utterance status of feedback, repetition as feedback could provide another refinement in attempting to measure language acquisition through feedback. Repetition, as a means for feedback giving and elicitation, is important for second language learners. We find repetitions of many different linguistic structures and they can have several functions (see ESF:Allwood and Ahlsén 1986; ESF:Vion and Mittner 1986). Repetition is a simple means of feedback giving for the learner who does not have many other means of expression. In this function, it is used by learners early in their acquisition. By adding a questioning intonation to the repetition, the learner also has a way of eliciting, for example, showing non-understanding or asking for clarification. All of these functions of repetition are probably acceptable in most languages, but they will be more or less common. Some learners start out with more attempts to use repetition than other learners, perhaps through source language influence. In a similar way, some learners will find more support for their use of repetition in the target language than others. The use of repetition in the different languages also has to be put in relation to the availability of other types of feedback in both source and target languages, as well as to factors like learner characteristics and activity type.

The use of repetition as feedback was studied in two ways. The total amount and share of repetition among the feedback units for the twenty learners in the three cycles was calculated and used as a basis for a general overview. Let us first have a look at the number and the relative share of repetitions in the feedback of the learners and the controls (see Table 8.5).

Table 8.5 shows a clear decrease iii the number of repetitions used for feedback from cycle 1 to cycle 3 for fourteen of the twenty learners and for seventeen for the twenty learners if reformulations are left out and pure repetitions alone or in combination with simple feedback is included (the second of the two C1-C3 columns in Table 8.5). This tendency is clear enough for it to be safe to assume that second language learners use repetition as an especially prominent type of feedback in early stages. This is also supported by the low shares of repetition for the controls.

Table 8.5 *Repetition - percentages in relation to total number of feedback units. Total number of FB units for each individual is given in brackets.*

			Cl	C2	C3	C1-C3 (incl. re-formulations)	C1-C3 (Pure repetitions excl. reformulation)
Sw	Fi	Man	12(77)	5 (94)	18 (155)	6	-6
		Leo	10(170)	5 (130)	6 (176)	-4	-4
	Sp	Nora	6(65)	4 (145)	5 (83)	-1	-4
		Fernando	7(71)	4 (97)	12 (200)	5	7
Fr	Sp	Berta	16(86)	14 (65)	17 (51)	1	-5
		Alfonso	36(55)	17 (82)	14 (56)	-22	-36
Ar	Zahra	16 (19.5)		13(261)	10 (220)	-6	-3
		Abdelmalek	63(139)	28 (109)	11 (67)	-42	-46
Du	Ar	Fatima	58(195)	5 (239)	5 (282)		-1
		Mohamed	10(246)	4 (299)	6 (155)	-4	-2
Tu	Ergün	13 (161)		17(151)	15 (193)	2	2
Mahmut	30 (252)			14(322)	13 (296)	-17	15
Ge	Tu	Çevdet	50(46)	17 (58)	6 (78)	-44	-37
		Ilhami	20(41)	2 (47)	6 (64)	-14	-14
	It	Marcello	69(49)	13 (84)	6 (83)	-63	-49
		Tino	50(67)	28 (67)	11 (112)	-39	-27
Eng	It	Lavinia	69(26)	14 (14)	2 (41)	-67	-32
		Andrea	43(21)	10 (59)	10 (79)	-33	-6
	Pu	Ravinder	3(69)	6 (89)	5 (40)	2	3
		Madan	33(86)	4 (70)	14 (123)	-19	-5
Mean relative share			31	11	10		

TL Controls: SL Controls:
Swedish   Adam 7(81)   Spanish Nora 7(91)
Eva   12(60)   Finnish Man 16(61)
English   Martin 9(34)
      Sheila 0(34)

Conclusions: Supplementing ML U with feedback derived measures
MLU gives a fairly indirect measure of linguistic proficiency. It therefore needs to be supplemented by other measures. In this section, we have suggested that some such measures could be derived from the study of interactional linguistic (narrow) feedback. The reason for this is that various aspects of feedback show a great sensitivity to the needs of naturalistic adult language acquisition. We have found the following correlations between feedback phenomena and language acquisition.

1. The relative share of feedback as measured through F'DU and FBW regularly decreases as a learner gets more proficient.
2. The relative share of singles, that is, utterances consisting only of feedback, and repetition used for feedback purposes decreases as the learner gets more proficient.

This means that an increase in MLU which is combined with a decrease in FBU and FM, a decrease in feedback singles and the use of repetition for feedback or a subset of these, is a more reliable measure of language acquisition and increase in linguistic proficiency than MLU alone.

However, in order to be really useful, what is said above should also be made sensitive to differences in activity demands. Activity must, therefore, be kept constant in using any of the measures under consideration as indicators of linguistic proficiency. The problem of activity variation will be examined in detail in Volume 11:11.2.

Thanks are due to two members of the Swedish team, Elisabeth Ahlsén and Joakim Nivre, for reading and commenting on this section.