This study investigated the effects of teacher feedback on changes in the spoken language performance of a group of advanced learners of English in the Netherlands. Participants were medical researchers attending a course entitled English for Medical Congresses. The course included a sequence of speaking tasks and culminated in a conference. Tasks were graded, and feedback was designed to increase in complexity. A combination of written feedback and audio- and videotape recordings were used. The study examined whether students became aware of, and mastered, language problems identified in teacher feedback as well as problems not identified in feedback, and whether they were aware of changes in their language performance. Analysis of data from student surveys, written feedback, language logs, and audio and video recordings indicated that students' perceptions of the value and effect of feedback were matched by real improvements in their spoken performance. The course outline, language log, and questionnaire are appended. (Contains 40 references.) (SM)
EFFECTS OF FEEDBACK ON PERFORMANCE:
A STUDY OF ADVANCED LEARNERS ON AN ESP SPEAKING COURSE

Tony Lynch and Joan Maclean (IALS)
EFFECTS OF FEEDBACK ON PERFORMANCE:  
A STUDY OF ADVANCED LEARNERS ON AN ESP SPEAKING COURSE 

Tony Lynch and Joan Maclean (IALS) 

Abstract 

Most language learners and teachers believe correction and other forms of feedback help improve L2 accuracy and fluency. This study set out to investigate the effects of feedback on changes in spoken performance among a group of advanced learners of English attending an intensive ESP course in the Netherlands. In this paper we assess the extent to which feedback from tutors helped highlight weaknesses in the participants’ spoken English and track changes in performance in those areas during the course. We conclude that the learners’ perceptions of the value and effect of feedback were matched by real improvements in their spoken performance. 

1. Introduction 

There is widespread, though not universal, agreement that language learners make better progress in the L2 if they receive overt feedback — though for a contrasting view, see Truscott (1996, 1999). Schachter (1983) went as far as to call feedback a ‘nutritional need’ for language learners, portraying it as having two dimensions — negative/positive on one axis and implicit/explicit on the other. In non-pedagogic interaction, feedback is more likely to take an implicit form, such as requests for clarification (implicit negative) and back-channelling (implicit positive), while in the conventional language classroom there is a tendency towards greater explicitness, for example, through formal correction (explicit negative) or praise (explicit positive). 

The giving and understanding of feedback is a complex process, and Carroll (1996) has made the point that, particularly in informal conversation, feedback is intrinsically inferential in nature. Even in the classroom, where teachers expect and are expected to provide feedback, learners’ responses to teacher feedback are bound to be ‘filtered’ through their individual interpretations of the teacher’s intention. In the 1970s, there were various attempts, stimulated by Corder’s work on the significance of learners’ errors (Corder 1967), to map the routes by which feedback on L2 errors is given and received in the language classroom (e.g. Allwright 1975, Chaudron 1977, and Long 1977). The resulting flowcharts show that the researcher’s task in representing interpretative
decisions is difficult enough; the task of the teacher and learner, to engage in the feedback process in real time, is even harder. In the next two sections we summarise recent work on feedback in those two domains - first from the technical literature of SLA research and then from the practical literature of classroom methodology (terms from Ellis 1997).

1.1 Technical literature

One of the developments in SLA research into feedback since the 1970s has been a shift from description to explanation - from categorising types of feedback to the investigation of what leads to successful uptake of feedback. This has applied to teachers' feedback on both written and spoken language, though here we will concentrate on the latter. (For recent reviews of research into feedback on L2 writing, see Grabe and Kaplan 1997, Hyland 2001, and Ferris 2002).

As we pointed out in the Introduction, there is no universal agreement among SLA researchers that feedback has any direct effect on learners' spoken performance; notably, Truscott (1999) has argued that oral correction by the teacher has no impact on spoken grammatical accuracy - following his similar claim that grammar teaching does not improve writing performance (Truscott 1996). However, the majority of SLA studies have borne out the assumption that some forms of feedback are effective in the short term, in the sense of leading to modification of the error and/or imitation of the correct form. One much-cited example is the study by Lyster and Ranta (1997) analysing the feedback provided by teachers in a Canadian primary school French immersion programme. Lyster and Ranta developed a descriptive framework combining categories from the negotiation of meaning and others from existing classroom observation schemes:

Table 1: A framework for analyzing feedback (Lyster and Ranta 1997)

<table>
<thead>
<tr>
<th>Explicit correction</th>
<th>provision of the correct form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recast</td>
<td>reformulation of all or part of the learner’s utterance</td>
</tr>
<tr>
<td>Clarification request</td>
<td>indication that what the learner has said is unclear or incorrect</td>
</tr>
<tr>
<td>Metalinguistic feedback</td>
<td>comments related to a problem in the learner’s utterance but without providing the correct form</td>
</tr>
<tr>
<td>Elicitation</td>
<td>a teacher may (1) elicit completion of their own utterance, (2) ask a question to elicit a correct form, or (3) occasionally ask a learner to reformulate their own utterance</td>
</tr>
<tr>
<td>Repetition</td>
<td>teacher’s repetition of a learner’s incorrect utterance, usually with marked intonation to highlight the error</td>
</tr>
</tbody>
</table>

Lyster and Ranta found that recasts accounted for more than half of all feedback, followed (in order of frequency) by elicitation, clarification requests, metalinguistic
feedback, explicit correction, and repetition. They argued that the predominance of recasts may well have been because the classes they studied were content-based, so that the teachers' and learners' attention was focused on the subject matter rather than the L2 form in which it was expressed. However, examination of the learners' responses to the six forms of teacher feedback showed that recasts were less effective than the other types in getting learners to produce the correct form. This bears out the findings of an earlier study of primary-age ESL schoolchildren (Oliver 1995), which found that the children did incorporate L2 points from recasts, but only in 10% of cases. Lyster and Ranta concluded that elicitation and metalinguistic feedback were most likely to succeed in producing correct forms from the learners.

Clearly, the issue of how teachers contribute to classroom discourse is important, since different pedagogic methods or approaches presumably require different patterns of participation and feedback. For example, in his proposal—or, rather, demand—for 'the Lexical Approach', Lewis (1993: 195) claimed that 'reformulation should be the natural response to learner error'. However, the evidence so far, from studies such as those of Lyster and Ranta, and Oliver, is that learners' uptake from implicit negative feedback is low.

Research within the framework of sociocultural theory (reviewed in Lantolf 2000 and Tarone 2000) has paid particular attention to the negotiation of feedback in collaborative dialogue between expert and novice, in Vygotskian terms. The outcome of one influential study (Aljaafreh and Lantolf 1994) was a model describing the graduated feedback available in dialogic interaction, shown below.

**Regulatory scale – Implicit (strategic) to Explicit**

0 Tutor asks the learner to read, find the errors, and correct them independently, prior to the tutorial.

1 Construction of a 'collaborative frame' prompted by the presence of the tutor as a potential dialogic partner.

2 Prompted or focused reading of the sentence that contains the error.

3 Tutor indicates that something may be wrong in a segment (e.g. sentence, clause, line) – "Is there anything wrong in this sentence?"

4 Tutor rejects unsuccessful attempts at recognising the error.

5 Tutor narrows down the location of the error (e.g. tutor repeats or points to the specific segment containing the error).

6 Tutor indicates the nature of the error, but does not identify the error (e.g. "There is something wrong with the tense marking here").

7 Tutor identifies the error ("You can’t use an auxiliary here").

8 Tutor rejects learner’s unsuccessful attempts at correcting the error.

9 Tutor provides clues to help the learner arrive at the correct form (e.g. "It is not really past but something that is still going on").
Aljaafreh and Lantolf’s scale was designed to categorise feedback on written L2 output, and would obviously need to be adapted for use in describing the options available in feedback on spoken output. Moreover, the one-to-one tutorial featured in their study is hardly typical of most language classrooms, and research is needed into more typical classroom settings.

Apart from the issues of the type and degree of feedback, a number of other aspects have come under scrutiny in recent research: the need to explore the distinction between ‘on the spot’ learning and ‘delayed learning’, and the particular problems of identifying the latter (Gass 2002); the importance of the learners’ perceptions of negative feedback, in the case of spoken language (Mackey, Gass and McDonough 2000) and writing (Leki 1991, Anderson, Benson and Lynch 2001); and the question – fundamental for the classroom – of why it is that a single exposure to feedback is sometimes enough for immediate uptake (Doughty 2001).

1.2 Practical literature

As is often the case, there is a striking difference between the assumptions of language teaching methodologists and those of SLA researchers. Authors who have discussed feedback from a pedagogic perspective seem to have been at least as much concerned with its social and affective repercussions as with its effect on individuals’ learning. Harmer, for example, calls for ‘gentle correction’ (2001: 107) and Rinvolucri is critical of teachers who provide ‘hamfisted feedback’ (1994: 288). Even the use of the word negative seems to be avoided in the methodological literature, which makes an interesting contrast with the unconcerned use of the term ‘explicit negative feedback’ in the technical literature.

The advice offered to inexperienced teachers tends to be presented in purely practical and simplistic terms, particularly in courses designed for initial teacher training. Here are some examples from methodology books dealing with feedback during group work:

‘Don’t correct unless the aim of the activity is controlled [and] unless a student asks you to... If the students need a lot of help and correction then the chances are that the task is inappropriate and/or beyond their capabilities’

(Gower and Walters 1983: 46)

‘Don’t correct mistakes. Make a note of anything serious and reteach it in another lesson’

(Byrne 1987: 79)
The teacher must sit or crouch down so that she or he is at the same height as the students'  
(Brown 1988, quoted in Wajnryb 1992)

The assumptions underlying these three prescriptions appear to be that: (1) there is a fixed maximum amount of correction that a learner can reasonably expect; (2) ‘anything serious’ must already have been covered in the course; and (3) standing up affects the acceptability or effectiveness of the feedback!

In the latest edition of perhaps the most widely-used British EFL teacher education text, Harmer (2001) follows Edge in dividing mistakes into slips, errors and attempts (Edge 1989) – the last category being cases where a learner tries to say something that is beyond their current L2 level. (These seem to be what Swain has referred to as holes in learners’ interlanguage). Harmer says the teacher’s main priority is errors, but does not explain how the teacher is supposed to establish in which category any given mistake belongs. His view is that

‘the feedback process is only finished once [the students] have made these changes. If students consult grammar books or dictionaries as a way of resolving some of the mistakes we have signalled for them, the feedback we have given has had a positive outcome’

(Harmer 2001: 112).

That is an odd statement, implying that feedback is completed, not by the learning of the correct form or item, but when the learner shows signs of having initiated progress towards it.

Increasingly, methodologists emphasise the need to extend feedback beyond correction of language form. Mendelsohn (1995) recommended a ‘triple focus’ in feedback in oral skills classes: linguistic, sociolinguistic (including paralinguistic features such as body language and gesture) and – not least – content. Similarly, Lynch and Anderson (1992) advised teachers to provide different types of feedback focus, depending on the speech genre being practised. In the case of information-gap tasks, they recommended a focus on the clarity and precision of the information; in role-plays, a sequential focus on overall strategy (cf. Mendelsohn’s ‘content’) first, followed by comments on the relative success of the communication, the learner’s use of the information provided, and finally language; and in the case of formal presentations, the use of an evaluation checklist covering all three of Mendelsohn’s areas.

2. The Study

2.1 Background

In a previous IALS research project we investigated the effects of simple practice on spoken performance (Lynch and Maclean 2000; Lynch and Maclean 2001). We found
that learners made language improvements over a cycle of classroom activity in which they repeated the same speaking task with different partners, without receiving any language feedback from the teacher between task episodes in the cycle. The actual improvements differed from learner to learner, but it appeared that the learners were each able to exploit the task in their own way to make at least short-term language gains.

These findings led us to speculate on the effect of teacher-initiated feedback in a sequence of speaking tasks. It is possible that feedback on points selected by the teacher could

1. interfere with the learner's natural processes and actually inhibit learning
2. help learners to notice the gap between their performance and the desired target, and so facilitate learning (Chaudron 1977; Schmidt and Frota 1986)
3. have no discernible effect, positive or negative

We were interested therefore, as a next step, to investigate the effects, if any, of feedback interpolated between speaking tasks in a sequence. The data source for our study was the English for Medical Congresses course that IALS has run in the Netherlands for more than ten years. It was chosen for three reasons. Firstly, feedback on spoken performance is central to its rationale and design. Secondly, the feedback procedures involve both written feedback records and audio- and video-recording of certain tasks, and therefore data collection would not interfere with or distort the normal course routines. Finally, the course caters for a similar learner community to that investigated in our earlier research into the effect of practice and repetition.

2.2 Description of the course

The version of English for Medical Congresses that we run in the Netherlands is a 25-hour course, lasting three and a half very full days. The course is sponsored by the Dutch Heart Foundation and the participants are medical researchers working in areas related to cardiovascular health, including health promotion. The majority are Dutch, in their thirties and engaged on PhD research. Their level of English ranges from early advanced to native-like.

The course consists of a sequence of speaking tasks, culminating in a conference on the final day of the course. The speaking task sequence is preceded, on Day 1, by language input sessions on grammar and pronunciation problems, collected from participants on previous courses. The points covered in these sessions can be seen as pre-casts (Samuda 2001), intended to raise awareness; they also represent a practical reference point and baseline, so that feedback can be given more quickly later in the course. The course then moves into a cycle of task performance and feedback, with one input session on Day 2 on phrasing and emphasis, and one on Day 3 on signalling language.

The tasks are designed to be graded, and the resulting feedback is also designed to increase in complexity. Day 1 features audio-recording of problem words; on Day 2, the
participants record 3-minute oral summaries, and on Day 3 group presentations are video-recorded in the morning, and rehearsals of individual conference presentations are video-recorded in the afternoon. Day 4 consists entirely of the final conference, at which each participant’s presentation is also video-recorded. After the course, the participants each receive a video of their own presentation, with written feedback. See Appendix 1 for a summary of the course activities and types of feedback.

The objectives of the course, from the designer’s and tutors’ point of view, are primarily English language improvement, with improvement of presentation skills a strong second objective. However, at the start of the course, when participants are asked to state their personal objectives, a few mention just language improvement, and a few note just presentation skills. Most include both, and feedback is therefore given on both aspects of spoken performance.

Of those who state language improvement as an objective, some prioritise accuracy and others fluency, according to their perception, of their own abilities. Fluency is obviously necessary for presentation skills, but these learners tend to rate both accuracy and fluency as important, for prestige among academic peers and for self-esteem. Increased complexity (Skehan 1998) is occasionally mentioned, mainly in terms of new vocabulary, but overall this has low priority for this particular group of learners, at least at the start of the course, compared to accuracy and fluency.

Continuous feedback records are kept throughout the course for each person; a sample is shown as Appendix 2. The groups rotate among the three tutors teaching the course, who each add to the continuous feedback record, which is kept by the learner. The learners also fill in Language Logs when preparing their rehearsal performance and final performance. As far as presentation skills are concerned, they complete self-evaluation forms of their own rehearsal, and peer evaluation forms for each other in the final performance. After the course, tutors analyse the video-recordings of the final presentations and send feedback on both presentation skills and English language.

2.3 Our research questions

The questions we set out to investigate in the study were these:

1. Is there evidence that English for Medical Congresses learners (a) become aware of and (b) master language problems identified in teacher feedback?

2. Is there evidence that they (a) become aware of and (b) master language problems not identified in teacher feedback?

3. Are the learners themselves aware of any changes in their language performance?

4. Do they have comments on what they find useful or not useful in the various modes of feedback on language improvement that they experience during the course?
2.4 Data collection

At the end of the course in November 2000, we asked the participants to complete questionnaires (see Appendix 3) on their perceptions of language improvement, and the value of the different types of feedback. We also asked their permission to make copies of their written feedback and recordings for research purposes. All 24 participants completed the research questionnaires, and 15 gave consent for the copying of feedback and recordings. Our data therefore consisted of:

A 24 completed questionnaires on perceptions of gain and value of feedback.

B 15 sets of:
- Feedback Forms and audio-recorded pronunciation feedback from tutors over three days: Two-minute talks, Informal talks, Using your voice; Reporting back, Data presentation, and Rehearsal. Also the final post-course feedback from the Conference on Day 4.
- Language Log 1 (Day 1) and Language Log 2 (Day 3).
- Audio-recordings of Two-Minute talks (Day 1), Reporting back (Day 2), Preparing the presentation (Day 3).
- Video-recordings of Rehearsal (Day 3), and final Conference (Day 4).

3. Findings

3.1 Questionnaire returns

So far we have analysed the 24 questionnaires, and the recordings and feedback records for 12 of the 15 who gave consent. We excluded materials from the remaining three participants, as they had not completed the Language Logs.

3.1.1 Perceived gains in language

Pronunciation gains were noted by 21 out of 24. Of these 21, 14 specified pronunciation of technical terms, and two specified phonological changes (from what we called "Dutch pronunciation").

Fluency gains were specified by only two out of 24, though two more mentioned increased confidence.

Gains in grammar were noted by eight out of 24 - some from feedback and some from the "pre-cast" session on grammar.

Although nobody had stated at the start of the course that increased complexity was their learning objective, four participants specified gains in the use of signalling phrases ("I didn't use them much before") and five others specified gains in the use of stress for
emphasising important words, and phrasing. Both these aspects had been introduced in input sessions and then practised in performance.

3.1.2 Perceived usefulness of performance, analysis and feedback

Overall, the highest ratings were given to

- viewing the rehearsal rated at 1 by 15 out of 24
- rehearsal itself rated at 1 by 13 out of 24
- tutor feedback on rehearsal rated at 1 by 13 out of 24

Language Log 1 activities (speaking with a partner, analysing recording oneself, discussing points from analysis with tutor) were given highest ratings by some, and lowest by others. It may be that individuals differed in the extent to which they preferred to be responsible for critiquing their own performance. Alternatively, it may be that some pairs worked better together as partners than others.

Interpretation is also made more difficult by the fact that not all of the participants did the Language Log 1 activity as instructed. Logistical problems with computers on the evening of Day 2 had created long queues to prepare Powerpoint slides, and some participants therefore recorded themselves alone, without speaking to a partner. We cannot tell from our data whether this was a factor in their giving the activity a high or low rating. Also, quite a number who had not done the task at all rated it as least useful; again, we cannot tell whether they found it least useful because they did not do it, or vice versa.

One enlightening finding, from the point of view of planning future courses, was that filling in Language Log 2 (while viewing the video-recorded rehearsal) received a low rating in comparison with viewing the video-recording and discussing it with the tutor. In retrospect, this is not surprising. In addition to its recording function, the Log was intended to encourage self-criticism/awareness, but such encouragement seemed unnecessary. They already analysed and evaluated as they watched their performance on tape (viz. the high ratings for the usefulness of viewing) and so completing Language Log 2 may well have been thought a purely bureaucratic requirement. The recording itself, however, remains useful as input to discussion with the tutor. So in EMC courses run since November 2000, Language Log 2 has been retained, but in a very much simpler format.

3.2 Some sample extracts from the data sets

For each learner we made up a datasheet, which traces the chronological course of various elements of speaking performance and feedback (including self-analysis in Language Logs) over the four days. In particular, we noted linguistic items – phonological, lexical or syntactic - that were highlighted in feedback and then occurred in subsequent performance. In such cases, we categorised the items in four ways:

1. "blithely wrong" - where the speaker appeared unaware of being wrong
2. wrong, but with evidence of attention and additional processing – e.g. hesitation, pausing and/or self-correction

3. right, but with evidence of attention and additional processing

4. right without apparent effort

Those four categories were marked in the datasheets as (1) **bold**, (2) *underlined*, (3) *italic* and (4) regular font.

We also compared and noted each individual's questionnaire responses, so that they could be compared with their performance and feedback records.

Finally we summarised what we felt were key points in the learner’s performance data and our interpretation of those. On the next two pages is a sample data sheet for Beryl; (like all other participants’ names in this report, ‘Beryl’ is a pseudonym).
<table>
<thead>
<tr>
<th>Pronunciation</th>
<th>Lexicogrammar</th>
<th>other</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pronunciation Introduction</strong></td>
<td>'through', not 'true'</td>
<td>Tendency to rising intonation at ends of sentences</td>
</tr>
<tr>
<td></td>
<td>gap (aspirate) junction, not 'gabjunction'</td>
<td></td>
</tr>
<tr>
<td><strong>Informal Talk: Feedback from tutor</strong></td>
<td>'potassium' like 'at'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>'hemi-channel', as /e/</td>
<td></td>
</tr>
<tr>
<td></td>
<td>'identified', as in 'eye'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>'hyperTROphic'; 'hyPERtrophy'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>'CHAracterised'; 'SYNthesis'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>'in vitro'; 'theory', not 't-'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I look (at) how...</td>
<td></td>
</tr>
<tr>
<td></td>
<td>'I may explain…’ should be ‘I can…'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>'put it into the mouse' = inject?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>To limitate = to limit</td>
<td></td>
</tr>
<tr>
<td><strong>Language Log 1</strong></td>
<td>Errors noted: Hypertrophy / hypertrophic</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Too little difference between ‘connexin' and ‘connexon’</td>
<td></td>
</tr>
<tr>
<td></td>
<td>‘construct’ (like ‘up’)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>‘through’, not ‘true’; ‘three’, not ‘tree’</td>
<td></td>
</tr>
<tr>
<td></td>
<td>‘microscopic’ (stress)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Errors noted:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>‘what I want to tell’ = ‘what I want to say’</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Query to tutor:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Can I say ‘The amount of marker protein indicates the transcription level’</td>
<td></td>
</tr>
<tr>
<td><strong>Presentation: Rehearsal</strong></td>
<td>gap junction (numerous)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>potassium</td>
<td></td>
</tr>
<tr>
<td></td>
<td>connexin (numerous)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>throughout; three - three</td>
<td></td>
</tr>
<tr>
<td></td>
<td>hypertrophy; hypertrophy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>hypertrophy (x 3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>synthesis; synthesis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>characterized</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Note to ‘change of accent’ for ‘determine’;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>‘characterized’; ‘hypertrophy’; ‘microscopy’</td>
<td></td>
</tr>
<tr>
<td><strong>Language Log 2</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Rehearsal: Feedback from tutor | is not known until now = it is still not known... | Nice visuals  
Stood well  
Good pace  
Good eye contact  
Clear loud voice  
Tend to end with very low fall |
|-------------------------------|-----------------------------------------------|------------------------------------------------|
| 'proteins' as /iːnz/  
'sinus/sino' as 'eye', not 'ee'  
eLEc'tron miCROscopy  
'axon/exon'  
'hypertrophy'  
'CHA'racterised'  
'mediated' as 'me...' | | |
| Presentation: Performance | is not known until now | Still high-rising intonation after clauses |
| | gap junction (numerous); gap junction (x2)  
sinoatrial; mediated;  
protein – protein – protein (x2) – protein (x2) – protein – protein – protein  
microscopy - microscopy  
hypertrophic  
hyperf'trophy - hypertrophy  
synthesis; construct  
characterized – characterized - characterized three; third | | |
| Questionnaire | Beryl says she noticed changes in language: pronunciation; trying to speak less monotonously, but 'during the presentation I kind of forgot this’. Usefulness grades: 1 – Tutor’s answers to Log 1 queries; rehearsal practice; viewing the rehearsal video; completing Log 2; tutor feedback on rehearsal; 2 – talking through with partner; analyzing recording using Log 1.  
“I think if I say something wrong now, at least I realize it, and the pronunciation of a lot of words I am doing better now – but still not everything, but I hope to improve this in the future, because I know now how to pronounce it correctly”. | | |
| **Summary** | Beryl’s final presentation very much confirms the summary she wrote in her questionnaire: some sounds/items are now correct; others are on the way to being correct; some are regularly incorrect. | | |
From Beryl’s datasheet we can trace, for example, the difficulty she had with syllable stress in the word *hypertrophy*. In the Informal Talk on Day 2 of the course, instead of stressing the second syllable, she produced the word with main stress on the first and secondary stress on the third. Our coding of occurrences of *hypertrophy* in the Pronunciation column of the datasheet on the previous pages shows that in nine occurrences from Informal Talk to final Presentation, Beryl improved the accuracy of the stress-placement after tutor feedback, as follows:

**Table 2: Stress-placement in ‘*hypertrophy*’ (Beryl)**

<table>
<thead>
<tr>
<th></th>
<th>wrong</th>
<th>wrong, after hesitation</th>
<th>right, after hesitation</th>
<th>right</th>
</tr>
</thead>
<tbody>
<tr>
<td>Informal Talk</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presentation:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rehearsal</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Note: Ticks represent occurrences in chronological order

Our data allow us to make a direct contrast of this particular feature between the overall improvement by Beryl and the lack of improvement by another participant, Fred, whose production the syllable stress in *hypertrophy* remained unhesitatingly inaccurate over 11 occurrences, despite his having received similar feedback from the tutor after the Day 2 Informal Talk.

**Table 3: Stress-placement in ‘*hypertrophy*’ (Fred)**

<table>
<thead>
<tr>
<th></th>
<th>wrong</th>
<th>wrong, after hesitation</th>
<th>right, after hesitation</th>
<th>right</th>
</tr>
</thead>
<tbody>
<tr>
<td>Informal Talk</td>
<td>(1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presentation:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rehearsal</td>
<td>(3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance</td>
<td>(7)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A particular difficulty for Dutch speakers of English is the fortis/lenis distinction, particularly at the end of a word that is followed by a word beginning with a voiced consonant. Examples in our data include gap junction, heard by the tutor as “gab junction”, and Heart Day, heard as “hard day”. Our data show that even over a relatively short period of course time learners can be made aware of such problems and can be seen (and heard) working towards greater accuracy. To illustrate this, we have data from Kate, a researcher working in health promotion. In talking about her work, Kate needed to refer to various awareness-raising activities, among which were Heart Days and a dance competition for school pupils. Table 3 shows her production of heart dance awards over the three days.

Table 4: Production of word-final /t/ in heart dance (Kate)

<table>
<thead>
<tr>
<th></th>
<th>wrong</th>
<th>wrong, after hesitation</th>
<th>right, after hesitation</th>
<th>right</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Informal Talk:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Tutor feedback</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>“I thought you said hard dance awards”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Presentation:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Rehearsal</em></td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td><strong>Presentation:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Performance</em></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

There we have some evidence that Kate became more accurate in her pronunciation and that after seeing her Presentation rehearsal and getting tutor feedback, she was able to get the expression right, first with some hesitation and then without, in her final Presentation.

Our illustrations so far have featured segmental and suprasegmental points, and while it is true that the majority of tutor feedback comments were about phonology, there were also cases of lexical and grammatical correction. One such came from Sue, who produced intermittent claudication (‘claudication’ = limping) as claudication intermittent. The way in which she increased the correctness of her word order is shown in Table 5.
Table 5: *claudication intermittent v. intermittent claudication* (‘Sue’)

<table>
<thead>
<tr>
<th></th>
<th>wrong</th>
<th>wrong, after hesitation</th>
<th>right, after hesitation</th>
<th>right</th>
</tr>
</thead>
<tbody>
<tr>
<td>Informal Talk: Tutor feedback</td>
<td>“claudication intermittent” = intermittent claudication</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presentation: Unpartnered rehearsal</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Presentation: Partnered rehearsal</td>
<td></td>
<td></td>
<td>(9)</td>
<td></td>
</tr>
<tr>
<td>Presentation: Performance</td>
<td></td>
<td></td>
<td>(7)</td>
<td></td>
</tr>
</tbody>
</table>

In this case, perhaps because decisions about word-order are under more conscious control than is the accuracy of segmental or supra-segmental production, Sue appears to have made a clear ‘switch’ after her first (unpartnered) rehearsal to 100% accuracy in the second rehearsal. All 16 instances of the target expression in the partnered rehearsal and final Presentation were correct and delivered without audible hesitation.

3.3 Overall summary of the data set

For reasons of space we will now summarise our findings in the form of the individual summaries that we made at the end of each participant’s data sheet, and will then tease out the main threads we have so far discerned in the data. The change of font indicates that these are working notes, made by the individual researcher, rather than a final analysis.

Ann
(Comment: her English was good in tests, and pronunciation is excellent. Yet she often said she felt her English was inadequate, in spite of reassurance from tutors and participants - see post-course feedback). She was hesitant and nervous in the Pronunciation Introduction, then too fast in Informal Talks, then steadier in pace. She
had episodes of processing difficulty (in Pronunciation Introduction) and in the Informal Talk. Unfortunately, only half the final Presentation was recorded, so we don't know how she managed at start. Overall she improved in managing her speed, intonation and hesitation noises. Faults: (1) overuse of Pres Cont and Pres Perf - no changes; (2) death doubled changed to rate of death instead of death rate. Her questionnaire responses are in line with our data, but she may not fully realise to what extent her episodes of processing difficulty are due to nerves rather than poor English.

Beryl
Beryl's final Presentation very much confirms what she said in her questionnaire: "I think if I say something wrong now, at least I realize it, and the pronunciation of a lot of words I am doing better now – but still not everything, but I hope to improve this in the future, because I know now how to pronounce it correctly". The data show that some sounds/items were correct by the final Presentation; others were on the way to being correct; and some were regularly incorrect.

Dave
A fluent and confident speaker overall. The most noticeable change in response to feedback is to his lip-rounding of up, etc. There is some evidence that he operates in two modes: in the final presentation he lip-rounded substances during question time, though he produced all the other instances in the prepared part of the Presentation correctly.

Ella
Ella wrote in her questionnaire 'I still try to improve my pronunciation, and before the presentation I planned trying not to speak so monotonous. But during the presentation I kind of forgot this, so I don't know whether it was (a bit) improved'.

As she said, Ella showed definite if not perfect improvement in pronunciation - mastering v/f in the key medical terms in her Presentation, and apparently more general v/f too. Note that blood was still wrong in the final Presentation, but correct in question time.

She corrected two out of her three lexicogrammar mistakes. She may have a strategy of avoidance: say/tell was not used after early errors.

Emma
Emma was a noticeably systematic user of feedback, going to the length of making her own notes on the tutors' comments on the feedback forms. She made excellent progress with the pronunciation of exacerbation, rationale, period, steroid, therapeutic; beginning to improve / study. I have the impression that all th sounds were more accurate. Study was self-corrected once in the final Presentation; but backslid on period in question time. No change in intonation. Exists of persisted in Presentation but was noted in the questionnaire as having been learnt.

Fred
What is not clear in the datasheet is that Fred had masses of Arabic L1 pronunciation and grammar problems - apparently randomly right and wrong. Labile vowels; tendency not to aspirate unvoiced consonants, etc. – e.g. pathway was recorded first as a th problem but later as p/b problem. Also, labile grammar with many mistakes in articles and sing/pl verbs - not much recorded by tutors.
His grading of feedback activities in the questionnaire suggests that he liked analysing himself with support and guidance... Perhaps gained in awareness of problems and perceived more change than actually happened.

**Frank**

Mastered points from feedback.

Pronunciation: *athero* and *ultrasound* were corrected in tutor feedback for *th* and *u* in the first two activities, then were OK in rehearsal and performance. *Coherence* was corrected in feedback to Informal Talk, and was then OK in rehearsal and performance.

**Lexicogrammar:**

1. *the Optical Coherence Tomography* was corrected (article deletion) after first activity, and OK in all following activities.
2. *research is carried out* in rehearsal is corrected to Continuous in performance
3. likewise to *sum up* is corrected to *to sum up so far*
4. *minimally invasive* correct with detectable processing difficulty (we had discussed this expression - he found it bizarre, and difficult to say).
5. *possibly avoided* *told* - using *pointed out* - but impossible to be sure.

**Jack**

Grammar was almost perfect, though I heard some minor errors not noted in tutor feedback. Pronunciation was also good in general, so his target points were all pronunciation of medical terms. Some appeared to be mastered, and others well under way to being mastered in final performance. Interestingly, *angiographic* was wrongly pronounced with /g/ every time, though he correctly used /dz/ instead of /g/ every time in *angiography* and *angioplasty*.

I thought he had taken advice of rehearsal feedback and was pausing slightly – and appropriately – when using signals.

**Kate**

Kate seems to have been able to direct attention to correcting some aspects of the pronunciation (e.g. *t/th* and *v/w*), but the combination of word-final *t* and word-initial *d* (*heart day, heart dance awards*) was trickier. On one occasion she did manage to produce *heart dance awards* successfully – which was the expression that the tutor had commented on having misunderstood. She was also able to self-correct the misplaced stress in *characteristic* and to get it right by slowing down.

**Nellie**

Nellie was able to act effectively on most of the points that had come up in feedback or Logs. In the case of *diagnosis*, some evidence that (like Dave) Nellie reverted to an incorrect version – /di-/ not /dai-/ - under the pressure of the question time section of the Presentation.

**Sue**

Interesting (unique?) use of unpartnered Rehearsal. She didn’t fill in Log 1, but we can see how she works on tutor feedback from earlier activities. Also interesting that this is without a partner... showing the value of practice even without feedback from a listener. She gets better at segmental *v*, too, and by the final presentation is producing a *v*-like sound in most cases.
Silvia

Silvia is apparently able to direct more effective real-time attention to phonemes than to intonation, even though she notes in Log 2 that she needs to improve the latter. That bears out what Dalton and Seidlhofer say about the relative difficulty of (self-)correcting intonation. Generally very accurate and proficient, which is perhaps why she regards only tutor feedback as 'very helpful'.

4. Discussion

Our earlier research (Lynch and Maclean 2000, 2001) focused on a similar group of English for Medical Congresses learners to those in this study, but there are important differences in the nature of the speaking tasks and feedback in the two studies. In the earlier research, the feedback was implicit, came from peers, arose in interactive discussion, and occurred in real time. In this present study, feedback was explicit (levels 10-12 on the Aljaafreh and Lantolf scale mentioned earlier), came mainly from the tutor, arose primarily in prepared presentations, and occurred after a time-delay to allow replay of the recordings.

Although we have not yet had time to analyse the data on a full case-study basis, we believe there is enough evidence to confirm that the widely expressed faith in tutor feedback does have an empirical basis, at least in the particular one-to-one form that the English for Medical Congresses course design allows. Most of the participants did improve their 'strike-rate' of correct L2 forms in the areas that had been brought to their attention by the tutor. Pending further analysis of our data, we would like to highlight aspects of participant behaviour and perception which we think are of wider interest and hope to investigate in greater detail.

The most striking finding was the learners' increased awareness of their language use, something that several commented on positively in their questionnaires. Throughout the course, the tutors emphasised that language learning involves both declarative and procedural knowledge, as opposed to the learning of scientific facts in the fields in which many of the participants specialise. A number of their questionnaire comments show that they recognised that improvements in English were likely to build up over time through a complex interaction of rule-learning and adequate opportunity for proceduralisation:

"I think if I say something wrong now, at least I realize it, and the pronunciation of a lot of words I am doing better now. But still not everything, but I hope to improve this in the future, because I know now how to pronounce it correctly".

(Beryl)

"I have learnt how to do it, now I have to practice!" (Ann)

The second point is related to the first. Our data show that learners do not reach a 'stable state' of complete accuracy in a particular language area; gains in accuracy during the series of speaking tasks are vulnerable to pressure, such as time (Skehan 1998). In particular, there is evidence from the datasheets that several learners performed at
different levels of accuracy in the prepared Presentation for the final day conference than in the question-time following the Presentation. Four participants (Dave, Emma, Fred and Nellie) made errors while dealing with audience questions in expressions that they had produced correctly in their Presentation. It would be understandable that when communicative stress (Brown and Yule 1983) rises, the speaker has less attentional capacity available for other aspects of performance. This is underlined by two responses to the questionnaire item What exactly did you change?:

“In language [I changed] some expressions – I also forgot one or two in the heat of the moment”. (Dave)

“Before the presentation I planned trying not to speak so monotonous. But during the presentation I kind of forgot this”. (Beryl)

On the other hand, there were two participants (Ella and Jack) who produced more accurate forms of words during question time than they had in the Presentation itself. Arguably, some individuals actually feel less stressed in a presentation once the prepared part is over and they are able interact with their audience more informally.

Of course, we would not claim that all changes in performance resulted from tutor feedback. Although our analysis has concentrated on tracing the sequential links between points made in written tutor feedback and the accuracy of those points in subsequent learner performance, there were many other inputs during the English for Medical Congresses course - the input lessons themselves (see Appendix 1), what we might call ‘collateral’ input from feedback given to other participants in the class, and the feedback that some received from peers in collaborative rehearsal.

The datasheets reflect a greater emphasis on phonology in tutors’ feedback notes than on grammar, and the words featured under Pronunciation are predominantly medical vocabulary. One reason for this is that individual learners cannot avoid the key terms in their academic area, which the course structure requires them to talk about informally, then formally, then in rehearsal, and finally in the conference. The second is that these are relatively advanced learners and make few syntactic and lexical errors in their specialist domain. We found very few instances in our data of repeated grammar error, or of corrected grammar error. In free performance, learners are not bound to specific syntactic structures; it could be that after feedback on grammatical mistakes, some learners avoided using the troublesome structure, thereby gaining in accuracy. But this can be only speculation.

5. Conclusion

Our study has gone beyond the very short-term focus of Lyster and Ranta (1997), which examined immediate post-feedback imitation by learners, and of Ohta (1995), which looked at learners’ use of recast expressions within the same lesson. Similarly, this research has extended the time-scale of our own previous study of the effects of practice.
(Lynch and Maclean 2000, 2001), involving a 20-minute cycle of speaking tasks, and has focused on potential effects during a task cycle lasting more than three days.

It is conventional in discussions of feedback in language learning to emphasise that feedback is a complex area — and it is. But our research so far suggests that if teachers provide feedback on spoken performance in written form and on an individual basis, and combine it with recordings of the performance, that may have the benefit of making feedback more accessible, more easily retrievable later and possibly more effective, especially in a series of linked pedagogic tasks such as those in the English for Medical Congresses course. The next stage of our study will expand this preliminary analysis into a fuller case-study description of the changes learners made to their spoken performances.

As always, the question arises as to how far the findings of an analysis of particular classroom data can be generalised to other settings. However, while it is true that the English for Medical Congresses course features an unusual degree of individualised feedback, it is by no means unique in doing so. At IALS, for example, a wide range of courses, both for general and specific-purpose language learners of English, include work on presentation skills, during which learners are given written feedback on their speaking. If it can be shown that a combination of written feedback and recordings helps learners to improve their subsequent performances, that will strengthen the argument that learners and teachers are right to regard the provision of feedback as a beneficial element of a speaking skills course.

References


### Appendix 1  Outline of course: activities and types of feedback

| Day 1 | 1 | Introduction and Opening activities. Quick tests, and meeting other participants through an introductory listening and speaking activity.  
  no feedback (relaxed activity) |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td><strong>Grammar and vocabulary: difficulties for Dutch speakers.</strong> Overview of errors most frequently made by participants on previous courses, including discussion of grammar items in test from previous session</td>
</tr>
</tbody>
</table>
|      | 3 | **Pronunciation difficulties for Dutch speakers.** Overview and practice, followed by Two minute talks. mini-presentations by participants on their own work.  
  2-min talks audio-recorded on personal cassette; personal written feedback from tutor on grammar and vocabulary, with pronunciation feedback on cassette; some points selected for review with group |
|      | 4 | Discussion skills. Asking and answering questions, giving opinions, commenting, agreeing and disagreeing; useful phrases, and practice.  
  no feedback, except informal according to tutor judgment (relaxed activity) |
|      | 5 | Editing abstracts.  
  individual review with tutors |
| Day 2 | 6 | In Informal talks. Participants talk interactively about their work, for about 15 minutes each (sitting round table, using white board as necessary).  
  15-min talks not recorded; personal feedback from tutor, and selected points for review with group, as in 3 above. |
|      | 7 | Using your voice. Using phrasing and emphasis for effective speaking: analytical exercises followed by 4-minute presentations, using OHP and graphs given by tutor.  
  4 min presentations at OHP not recorded; feedback as in 3 and 6 but including comment on use of voice, stance, eye contact, etc. |
|      | 8 | Reporting back. Using audiocassette input, each participant prepares a 3-minute oral summary (to report sitting round table).  
  3-min oral summary audio-recorded on personal cassette, feedback as in 3 and 6 |
|      | 9 | Preparation of conference presentations. Individually drafting outline and Visual aids; talking informally through outline and visual aids with a partner; audio-recorded, then noting corrections and queries, and consulting tutor.  
  Language Log 1: i.e. participant notes corrections and queries from discussion with partner, and then from tutor |
| Day 3 | 10 | Presentation styles and techniques. Consideration of video-recorded examples of presentation techniques, with particular focus on language.  
  Blank Speaker Evaluation Form given out for use in final conference, and to focus practice before it, (form more on presentation than language skills) |
|      | 11 | Data presentation. Participants work together on adapting tables of data for visual presentation, and then practise style and technique in 5-minute group presentations.  
  5-min group presentations video-recorded and then reviewed immediately; peer discussion with tutor: written feedback from tutor, emphasis on presentation skills |
|      | 12 | Rehearsal of conference presentations. Individually, with tutor (while others prepare posters).  
  10-min presentation recorded on personal video (also audio if liked). Viewed later by participant with a partner, using Speaker Evaluation Form, also Language Log 2 with corrections and queries; then discussed with tutor, who gives written feedback on language and presentation skills. |
|      | 13 | Poster session. Practice in asking and answering questions quickly and concisely.  
  no language feedback (relaxed activity) unless for something striking; group evaluation of posters |
| Day 4 | 14 | Conference. 10 minute papers followed by 3 minute question-time  
  Video-recorded. Peer evaluation through Speaker Evaluation forms, written after each presentation, given to speakers at end of conference. Videos reviewed post-course by tutors and sent to participants with written feedback on language and presentation skills. |
Appendix 2  LANGUAGE LOG : preparing your presentation

1  Reading the abstract: pronunciation feedback from tutor (Monday evening)

2  Talking through your presentation to a partner (± audio-recording): note any mistakes and corrections and bring queries to a tutor (Tuesday evening)

   WRONG or DOUBTFUL   YOUR CORRECTION   FEEDBACK FROM TUTOR

   if there is not enough space please use the other side of this paper

3  Reviewing your video-recorded rehearsal: first fill in speaker evaluation sheet with a partner; then note below language mistakes; then review with tutor for feedback (Wednesday evening)

   WRONG or DOUBTFUL   YOUR CORRECTION

   if there is not enough space please use the other side of this paper
Appendix 3

Research questionnaire

This course was designed to help you improve both language and presentation skills. Using video and feedback is a well-established way of improving presentation skills, but our research interest is in the development of language skills. We would very much like to have your comments on the questions below.

If you have already agreed to contribute feedback and recording data to the project, please sign here again (so that we can relate this questionnaire to your data)

If you have not agreed, there is no need to sign, but we would be pleased if you would still fill in this questionnaire.

1.1 In your conference presentation this morning, did you notice any changes, compared with your rehearsal in content? YES / NO please circle

in language? YES / NO please circle

1.2 If so, what exactly did you change?

1.3 For each of these changes, could you indicate whether it was planned or unplanned?
Please put P or U against each “item” above.
2 Grade the following for how much they helped you from the point of view of language improvement, using this 1-4 scale:

1 very helpful
2 helpful
3 not very helpful
4 no help at all

- talking through the presentation with your partner (planning stage)
- analysing the audiorecording yourself (Language Log 1)
- getting answers to Language Log 1 from a tutor
- the practice of the rehearsal itself
- viewing the rehearsal video
- completing Language Log 2 (on the rehearsal video)
- feedback from a tutor on the rehearsal video
- other (please specify)

3 Do you think you have made any language improvement during this course? If so, please note here anything you have learned in pronunciation, vocabulary, grammar, or use of language. Please be as specific as possible.

Thank you very much for your help.
NOTICE

Reproduction Basis

X This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").