

BNCM13
BNC Task Group Managers
Minutes for meeting of 10th December, 1991

Dominic Dunlop

13th December, 1991

Present:

MB	Michael Bryant	Lancaster
GB	Gavin Burnage	OUCS — observer
LB	Lou Burnard	OUCS — observer
JC	Jeremy Clear	OUP
SC	Steve Crowdy	Longman
DD	Dominic Dunlop	OUCS

Opening of meeting. The meeting convened at 15:00. DD agreed to write the minutes; JC agreed to chair the meeting.

1 JHC's revised project milestones and schedule

The group discussed TGCW14¹ BNC *work programme: revised schedule and milestones*, which had been circulated by JC. A number of minor amendments were suggested:

- Pending agreement by the Project Committee on the manner of Chambers' participation in the project, Chambers should not be shown as responsible for any activity.
- The *draft written design spec.* activity should be shown as complete, not pending.
- The *tag set spec.* activity should not be shown as complete: input received from the SALT list as a result of the canvassing of the current version of the tagset needs to be acted upon. A new version of TGDW01, *Proposal for basic grammatical tagset*, should quickly be generated and circulated with a view to discussion at a task group D meeting, to be called as soon as possible by MB. MB
- The *collect spoken text* activity should be subdivided into two tasks: *data capture to agreed spec.* and *convert to CDIF*. Longman is to be responsible for the first of these; OUCS for the second.
- It is not clear what is involved in the *ISER prototype* and *ISER final* activities, so this part of the schedule may be subject to change. (See next section.)

The document was otherwise acceptable. JC will circulate a minor revision, and will forward the document to the DTI. JC

¹I have assigned "project-wide" document numbers to the project plan and to these minutes, there being no filing category for "task-group-wide" documents. If anybody objects, please let me know.

2 “Input Search and Retrieval” software

A Lancaster-based activity associated with the production of “input search and retrieval” software was included in the original BNC project proposal, but there is no specification of the intent of, scope of, or level of finish expected from, the activity. The group agreed that it was separate from the activity of recording and disseminating Corpus materials, which is the responsibility of OUCS. OUCS is also responsible for monitoring copyright issues once electronic texts are published. There was a brief discussion of what this might entail.

MB stated that he presumed that the tools in question would be of use in utilizing the Corpus data after users had obtained it from OUCS. Potentially, a great deal of effort could be expended in the production of an all-encompassing package of commercial-quality software. This was clearly beyond the scope — and budget — of the BNC project. GB, citing BNCR12, pp 7–8, said that he had identified a clear need for software tools when he visited potential users of the Corpus in Ireland. LB stressed the importance of taking advantage of the Corpus’ SGML encoding in any tools distributed by the project, and that the tools should not focus solely on the needs of lexicographers.

JC distributed BNCW15, *Software tools*², which sets out possible areas for Corpus-related software development. LB informed the group, that, under academic licencing terms worked out between OUCS and Software Exoterica, OUCS would be using XGML and related products in the development of tools, which would subsequently be put into the public domain. There was some discussion of the possibility of the production by Longman and OUP of Corpus-related tools which might be made freely or commercially available. (Both companies have commercial software arms.) However, both foresaw only the creation of in-house programs which would fill their own requirements, but which would not be made available to others inside or outside the project.

Ultimately, it was decided that MB should produce a proposal on the topic of IS&R for discussion at the Project Committee meeting of 14th January, 1992. The proposal should be circulated to task group managers for review and comment before that date if at all possible. The group agreed that it was inappropriate to refer the matter to the Advisory Committee at this stage. **MB**

3 Any other business

LB informed those present that he had made an approach to Reuters with a view to obtaining newswire material for inclusion in the Corpus. The group would be kept informed of further developments³.

There being no other business, the meeting closed at 17:00.

²If JC or others are against numbering this document, let me know. As I see it, if MB refers to it in his report, it needs a number, otherwise it can vanish from the record.

³A letter subsequently received from Reuters suggests that the company is willing to provide material after mid-1992.