



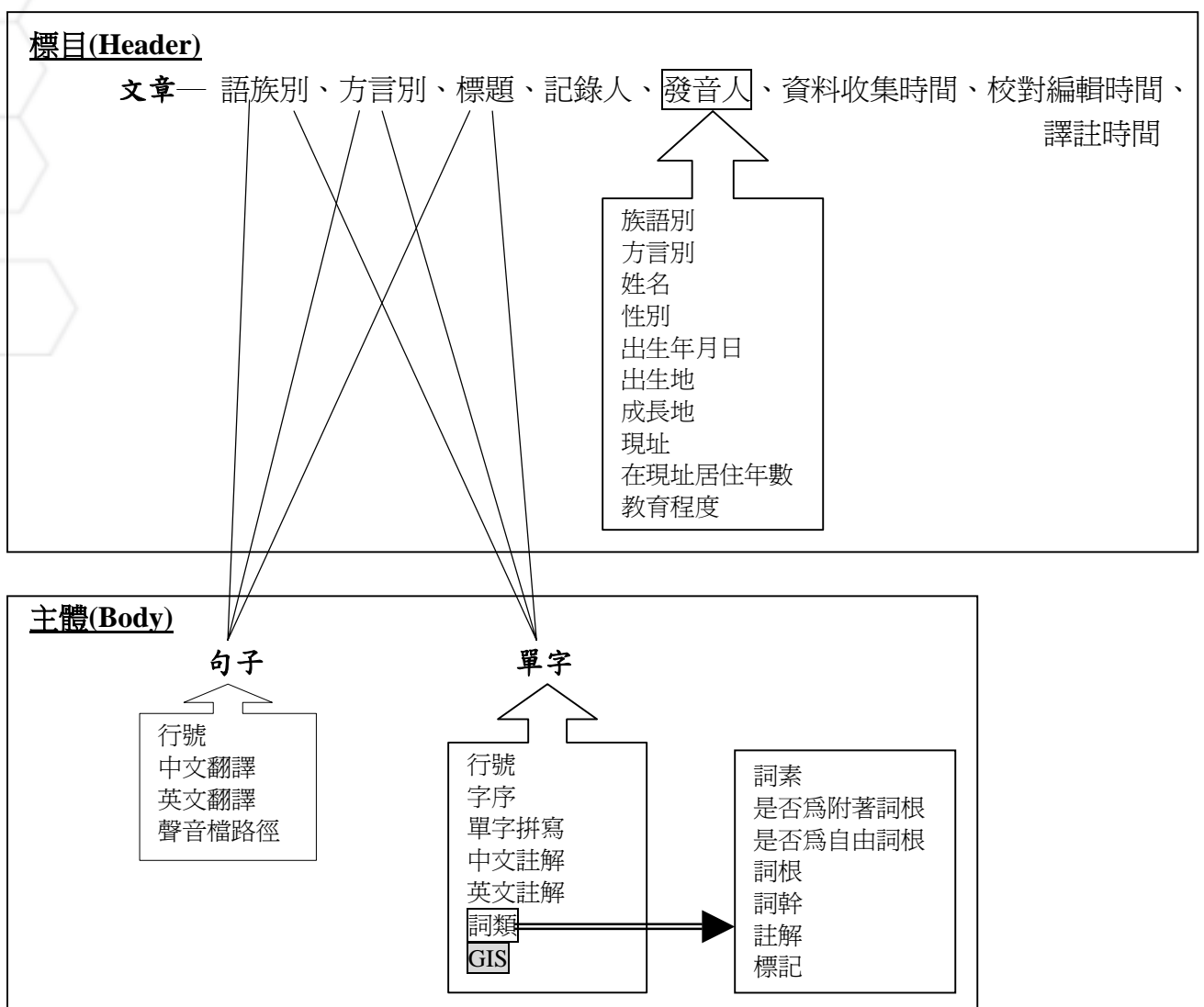
「台灣南島語數位典藏」後設資料元素分析報告(一)

數位典藏國家型科技計畫 後設資料工作組 研製

- 時間：91年6月17日
- 地點：語言所317室
- 參與人員：台灣南島語數位典藏計畫—齊莉莎老師、余清華
後設資料工作組—龔家珍、翁翠霞

一、「台灣南島語數位典藏」後設資料欄位初步分析

- 「台灣南島語數位典藏」欄位分析：
 - (1) 具有描述語料本身(標目 header)描述語料內容(主體 body)兩類欄位
 - (2) 描述語料內容欄位可分有兩種層次：句子、單字
 - (3) 同源詞與非同源詞的 GIS 查詢(未提供欄位)
- 欄位的結構層次為：





二、國際上有關語料庫的後設資料標準

- 描述語料文件標目的後設資料標準
 - (1) **OLAC**—根據 OAI 架構爲了資源共享和發現資源的目的所建立的後設資料標準，欄位名稱主要依據 DC
 - (2) **IMDI Metadata Elements for Catalogue Descriptions**—主要描述已出版的語料集(published corpora)，以作爲頂層的目錄
 - (3) **IMDI: Metadata Elements for Session Descriptions**—描述的目標是針對多種模式、多種媒介的語料資源(multi-modal multimedia language resources)。URL:
http://www.mpi.nl/world/ISLE/documents/docs_frame.html
 - (4) **OLIF (Open Lexicon Interchange Format) v.2 administrative data categories**—爲了交換詞彙資料和術語而建立的媒介。URL:
<http://www.olif.net/>
 - (5) **AILLA (The Archive of the Indigenous Languages of Latin America)**—爲描述原住民語言語料庫的標目資料。URL:
<http://www.ailla.org/pc/mainindex.html>
- 描述語料內容本體的後設資料標準
 - (1) **IMDI: Metadata Elements for Lexicon Description**—描述目標爲多語的詞彙(lexical headword)及其特徵。URL:
http://www.mpi.nl/world/ISLE/documents/docs_frame.html
 - (2) **OLIF (Open Lexicon Interchange Format) v.2 The Structure and Content of the Body**—爲了交換詞彙資料和術語而建立的媒介。URL: <http://www.olif.net/>
- 同一後設資料標準具有描述語料文件標目和語料內容本體的標準
 - (1) **TEI (Text Encoding Initiative)**—爲用於電子形式交換的文本編碼標準。URL: <http://www.tei-c.org/>

三、「台灣南島語數位典藏」欄位疑問

1. 元素「文章」內的次元素「發音人」是否和「發音人」元素所描述的對象是相同的？
2. 同一份後設資料記錄是記錄同一文章或同一次錄音的句子和單字？還是可以不同文章或錄音的單字或句子？
3. 如果最小的藏品單位是「單字」，而「單字」需要一筆後設資料記錄。但是同一個單字可能出現在不同的句子、文章，或由不同發音人所說出，這樣是否會造成「句子」和「文章」的紀錄要無限地重複著錄？



4. 如果「單字」和「句子」皆為一筆記錄，是否需要建立兩套後設資料元素需求表單？
5. 在 06 著錄範例、07 系統屬性功能需求、08 系統查詢功能需求均無提供「詞類」欄位的相關範例或功能，能否提供以作為分析和系統設計的依據？

四、會議結論

1. 原南島語計畫 Metadata 元素描述的對象是「發音人」，經討論後決定以「文本」為描述對象；本年對先進行語料標目(header)後設資料的建立。齊老師會在 6/21 把更正過的欄位回傳給後設資料工作組。(已於 6/19 寄回後設資料工作組)
2. 有關發音人、翻譯人、校對人的資料建議可建立人名權威檔
3. 本年度不進行同源詞與非同源詞 GIS 查詢系統
4. 南島語計畫、語言典藏計畫及後設資料工作組將擇期討論有關語言典藏的後設資料，以及南島語計畫如何和 OLAC 標準相連結等問題

五、後設資料工作組須提供給南島語計畫的資料

1. 人名權威檔參考規範

六、南島語計畫須提供給後設資料工作組的資料

1. Metadata 需求欄位更新表單

七、南島語計畫對後設資料系統的疑問

Q: 由於南島語計畫已利用 Access 建立一資料庫，如果新系統建立完成，能否將舊資料庫裡的資料整批匯入，而不要單筆輸入？