

目錄

通用轉錄指導方針	2
字母的大小寫	2
字母詞、首字母縮略語和拼出來的字母	2
標點符號的用法	2
數字	3
特殊符號	3
已命名的實體	4
數字符號	4
日常使用的非正式英語	5
不合語法的單字	5
語言模糊	5
• 同音異義詞	5
• 變體拼寫	5
• 專有名詞	6
區域拼寫差異	6
常用外來語	6
英語的重音符和變音符	6
側面搜索 / 上下文音頻	6
有效標籤	7
• <UNKNOWN/>	7
不可理解的說話聲音	7
外語	7
方言	7
截斷的說話聲音	8
錯誤發音	8
• <FILL/>	8
不完整的單字	8
填補詞	8
感嘆語和驚嘆詞	9
重複	9
• <CNOISE/>	9
唱歌	10
• <SN/>	10
單獨出現的突發噪音	10
重疊性突發噪音	11
多個說話者和說話聲音重疊的標籤	12
• <OVERLAP> <OVERLAP/>	12
其他資源	14
• 辭典	14

通用轉錄指導方針

字母的大小寫

如果遇到中英文同時出現在一句話的情況，請在中文和英文之前加上一個半角空格。除了字母詞和拼出來的字母外，所有的英語詞彙都必須用小寫轉錄。

例子：

- 你知道 **maria** 去了美國的消息嗎
- 你會用 **google chrome** 嗎

備註：轉錄員需要根據說話者的實際發音進行正確的轉錄。如音頻中的說話者說的是「嗨」，則應該轉錄成中文，而不是英文的「**hi**」。

字母詞、首字母縮略語和拼出來的字母

只有**字母詞**和**拼出來的字母**才應該大寫。

當一個字母被拼出來時，必須先加一個半角空格，然後用大寫字母書寫，再在後面加上一個半角句號，最後再加一個半角空格。

例子：

- 圖像儲存在我的 **U. S. B.** 上
- 正確的拼寫是 **A. M. Y.**

在轉錄時，如果遇到了由「大寫字母和連字號」組成的詞彙，例如「**T-shirt** 和 **X-ray**」，請按照上述拼寫規則刪掉連字號，轉錄成「**T. shirt** 和 **X. ray**」。

同理需轉錄成 **O. K.**，而不是 **okay**。如果說話者只發音 **K.**，那麼就只轉錄成 **K.**。

至於 **Z.**，不管說話者發音是「**zee** 或 **zed**」，都需轉錄成 **Z.**。

至於**首字母縮略語**（可以作為一個普通單字直接念出來的詞彙），如「**nato** 或 **fifa**」，則需要以小寫轉錄。

標點符號的用法

除**所有格撇號**和**連字號**外，**不可**使用任何標點符號，包括句號、逗號、分號、冒號，感嘆號和問號等。

例子：

- **錯誤** - 如果能請假的話，我週四會開車去 **sainsbury's**。那天的天氣應該很好！
- **正確** - 如果能請假的話我週四會開車去 **sainsbury's** 那天的天氣應該會很好

口頭標點符號，例如「**google dot com**」等，應該遵循「**聽到什麼轉錄什麼**」的規則進行轉錄，而不是直接使用標點符號。

例子：

說話者**口頭說**：“你輸入 yahoo dot com，然後電郵至 smith 小老鼠 yahoo dot com。”

- **正確**：你輸入 yahoo **dot** com 然後電郵至 smith **小老鼠** yahoo **dot** com
- **錯誤**：你輸入 yahoo.**com** 然後電郵至 smith **@yahoo.com**
- **錯誤**：你輸入 yahoo **dot** com，然後電郵至 smith **at** yahoo **dot** com
- **錯誤**：你輸入 yahoo **點** com 然後電郵至 smith **艾特** yahoo **點** com

備註：如果說話者明確用英語說「**at**」，則應該遵循「**聽到什麼轉錄什麼**」的規則，寫成「**at**」，而不是「**艾特**」。

如果說話者明確用中文說「**艾特**」，則應該遵循「**聽到什麼轉錄什麼**」的規則，寫成「**艾特**」，而不是「**at**」。

備註：請多多注意**帶有連字號**的英語詞彙。如果連字號在該單字中具有**語義價值**，則需要轉錄出來。

如果該英語詞彙在刪除了連字號之後不被 Microsoft Word Speller 所識別，例如「**openfaced**」，則應在轉錄中使用連字號。

以下是**應該使用連字號**的例子。

例子：

- 我們去體驗星巴克 **drive-thru** 服務吧
- 他現在是全聯的 **part-time** 店員

數字

所有數字都必須轉錄成漢字。

例子：

- 我出生於**一九九二年六月三日**
- 她吃了**五碗**牛肉麵

特殊符號

轉錄員應按照說話者**所說的那樣**進行轉錄，始終遵循「**聽到什麼轉錄什麼**」的規則。轉錄中不應該出現無法念出來的**特殊符號**。

以下是**錯誤**的轉錄例子：

- 咖啡 **&** 茶
- 你是 **#1**

以下是**正確**的轉錄例子：

- 咖啡**和**茶
- 你是**第一名**

備註：轉錄員需要根據說話者的**實際發音**進行正確的轉錄。

如說話者說的是英語的「**and**」，則應該按照**實際發音**轉錄成英文。

如說話者說的是中文的「**和**」，則應該按照**實際發音**轉錄成中文。

以下是三個最為常見的特殊符號以及與之對應的可能讀法與寫法。

@	at 符號、艾特、小老鼠、at sign、at
&	和、兼、及、ampersand、and
#	數字符、井號、number sign, pound sign, hashtag

備註：以上所列出的「與特殊符號對應的可能讀音與寫法」只是其中幾個最常用的讀法與寫法。如果說話者**使用其他讀法**，轉錄員應當遵循「**聽到什麼轉錄什麼**」的規則進行轉錄，將特殊符號轉錄出來。

已命名的實體

人名、公司名稱、歌曲名稱、國家名稱、應用程式名稱等都應該按照上述規則轉錄（例如，使用小寫）。受版權保護的條款（如商標、品牌名稱、註冊名稱）也應以同樣的方式處理。名人、公司、歌曲、應用程式等都應該使用它們已知的商標或官方拼寫進行轉錄。

如果不確定如何準確地拼寫這些已命名的實體，請通過網路搜索正確的拼寫。

例子：

- **OneDrive** → **onedrive**
- **McDonald's** → **mcdonald's**
- **Coca-Cola** → **coca-cola**
- **Ice-T** → **ice T.**

如果不能確定正確的拼寫，請根據上下文做出最好的判斷。

數字符號

不可使用數字符號，必須使用與該符號相關聯的單字進行轉錄。

例子：

- $2 + 2 = 4$ → **二加二等於四**
- $6 \times 3 = 18$ → **六乘以三等於十八**

日常使用的非正式英語

如果說話者清楚地說出了被正式承認為單字的非正式英語，那麼就轉錄非正式英語。

例子：

- i **wanna** know 你到底行不行
- you **gotta** show 別人你才是冠軍

如果非正式英語沒有被正式承認為單字，那麼就要使用正式的官方拼寫。請不要自行創建新詞彙。

例子：

- **full of** → **fulla**
- **like a** → **lika**

如果說話者明確地說了「**哈哈**」（這與真正的笑聲不同，真正的笑聲應歸類成噪音）、「**隆隆**」、「**砰**」等等，並且辭典確實**正式承認**它為一個單字的話，請轉錄它們。

不合語法的單字

對於在對話內容中**不合語法**但清晰表達出來的單字，請按口語**所說的那樣**轉錄下來。

以下是**正確**轉錄的例子：

- 在這裡你會找**了**所有證據
- 她坐在最**終**一排座位上

備註：以上句子在指定內容中**並不符合語法規則**。但是，如果說話者確實說出了這些單字，那麼這些文字將被視為**正確的轉錄**。

語言模糊

語言模糊可以分為幾種形式：

同音異義詞（發音相同但拼寫不同的單字）

例子：

- 報復 / 抱負
- 行李 / 行禮

變體拼寫（具有多個拼寫的詞彙）

例子：

- 外形 / 外型
- 精彩 / 精采

專有名詞

例子：

- 嘉儀 / 佳怡
- 李成 / 李辰

對於同音異義詞、變體拼寫和屈折變化，只需轉錄你認為最有可能出現在該內容中的單字即可。

對於專有名詞，你可以使用最常用或最常見的變體。

區域拼寫差異

請注意遵守所轉錄語言的正確拼寫規定。在引用專有名詞時可能會出現例外情況。例如，如果正在轉錄國語，你需要轉錄「[國立臺灣大學](#)」，因為這是官方寫法。

常用外來語

如果外來語/外來字是該語言常規詞彙表的一部分（即大多數說話者都能夠理解的詞彙），請使用外來字轉錄，例如，可以轉錄「[ciao](#) 或 [bonjour](#)」。這些都是大多數說話者能夠接受和理解的詞彙。

英語的重音符和變音符

遇到特定人物、地方、組織/機構、商標以及在辭典中受到普遍認可的英語詞彙，請使用重音符或變音符。

例子：

- Beyoncé → Beyoncé
- San Sebastián → san sebastián
- Citroën → citroën
- Soufflé → soufflé

側面搜索 / 上下文音頻

使用側面搜索有助於理解上下文或識別地點、人物、商標等名稱。網路應用 (webapp) 中提供的上下文音頻也能大大提高對轉錄音頻的理解。周圍音頻則可以幫助確定單字是在音頻的開頭還是結尾被截斷。如果不確定的話，請檢查周圍音頻。

有效標籤

<UNKNOWN/>

將在以下情況中使用 <UNKNOWN/> 標籤：

不可理解的說話聲音

當聽到明顯的人類言語，但無法理解一個或多個實際單字時，請使用此標籤。

導致出現不可理解的說話聲音的原因可能是口音很重、語速過快、發音差、音質差或其他任何原因。

如果遇到在某些音頻中，部分的說話聲音是可理解的，而有一部分是不可理解的話，請對不可理解的部分使用 <UNKNOWN/> 標籤，並繼續照常轉錄可理解的部分。

外語

- 部分說話聲音為外語

如果一段話中的部分說話聲音為外語，請將該部分標記為 <UNKNOWN/>，但照常轉錄其他單字。

備註：如果說話者使用英語說出特殊符號，例如「and」，轉錄員不應該將該單字標記為 <UNKNOWN/>。請根據音頻的實際狀況作出判斷，或詢問負責人加以確認。

- 全部說話聲音為外語

如果一段話中的全部說話聲音為外語，請將整段話標記為 <UNKNOWN/>。但是，如果說話者提到一個以外語命名的實體，請轉錄該實體。

方言

- 說話聲音為聽得懂的方言

如果一段話中的說話聲音為普通台語，請照常轉錄該方言，不要標記為 <UNKNOWN/>。

例子：

- 伊將青春嫁置灶跤 → 她將青春嫁在廚房

- 說話聲音為聽不懂的方言

如果一段話中出現了少量聽不懂的方言，請將該部分標記為 <UNKNOWN/>，但照常轉錄其他單字。

如果一段話中出現了大量聽不懂的方言，則應放棄該音頻並聯繫負責人。

備註：台灣轉錄員應當可以正常轉錄普通台語。

截斷的說話聲音

當一段話中的第一個單字和/或最後一個單字在錄製的開頭或結尾被切斷（**截斷**）時，請使用 **<UNKNOWN/>** 標籤。

備註：如果整段話都是 **<UNKNOWN/>**（不可理解或是外語），請不要添加任何附加標籤。只需使用單個 **<UNKNOWN/>** 標籤來標記整段話即可。

備註：當說話聲音是清晰的，但因為各種原因（例如同音異義詞、變體拼寫等）而變得模稜兩可時，請勿使用此標籤。

備註：不要使用連續（不間斷）的 **<UNKNOWN/>** 標籤。如果說話者連續說了幾個不可理解的單字，只需要使用一個 **<UNKNOWN/>** 標籤。

錯誤發音

對於**意思不明**的錯誤發音，請使用 **<UNKNOWN/>** 標籤。

對於**意思明確**的錯誤發音，請使用標準的詞彙。

例子：

- **羅**杉磯在美國加州（念錯了「ㄌㄨㄛˊ」的聲調）→ **洛**杉磯在美國加州
- 暴徒的行為太可**餓**了（可以從上下文推斷出真正的意思）→ 暴徒的行為太可**惡**了

<FILL/>

當出現以下情況，就必須應用 **<FILL/>** 標籤：

- (1) 未完全說出的單字，
- (2) 缺乏任何語義價值的話語。

不完整的單字

當單字被“破壞”且說話者沒有完整說出單字時，請使用 **<FILL/>** 標籤。

例子：

- 我要點漢堡和 **ㄩ**... 漢堡就可以了 → 我要點漢堡和 **<FILL/>** 漢堡就可以了
- 明天一定要給我打 **ㄉㄨㄟ**... 電話聽到沒 → 明天一定要給我打 **<FILL/>** 電話聽到沒

填補詞

如果說話聲音中包含沒有語義價值的聲音，例如呃、嗯、啊等填補詞，請使用 **<FILL/>** 標籤。

例子：

- **呃**你晚餐想吃什麼 → **<FILL/>** 你晚餐想吃什麼
- 也許我們**嗯**沒有**嗯**準備好吧 → 也許我們 **<FILL/>** 沒有 **<FILL/>** 準備好吧

備註：有時「呃、嗯、啊」等單字是具有語義價值的。轉錄員應根據實際情況做出判斷，看看該單字是否屬於填補詞。

備註：如果說話者使用正式的單字作為填補詞（例如「就像」、「我是說」），請轉錄這些詞彙。

備註：連續說出填補詞的話，只需使用單個 <FILL/> 標籤。

例子：

- 這位**呃**先生這邊請 → 這位 <FILL/> 先生這邊請

感嘆語和驚嘆詞

感嘆語和驚嘆詞（例如**嗯哼**和**嗯**）被視為正常詞彙，而不是填補詞，因此需要轉錄它們。它們確實具有語義價值，因此有別於諸如**呃**、**嗯**、**啊**等填補詞。

例子：

- 說話者 1 - 你要走了嗎 → 你要走了嗎
- 說話者 2 - **嗯**我要走了**呃**下次再聊吧 → **嗯**我要走了 <FILL/>下次再聊吧

在以上例子中，**嗯**是用於表示同意的感嘆語。相反，**呃**只是口頭填補詞，所以應使用 <FILL/> 標籤來標記該部分。

重複

如果單字沒有被破壞，則不應該使用任何標籤標記**重複**的話語。不管說話者重複了多少遍，都只需轉錄說話者所說的每個單字即可。

例子：

- 你你這也太太過分了 → **你**你這也太太過分了
- 君子動動口不動手 → 君子**動動**口不動手

<CNOISE/>

如果在音頻的大部分時間中都出現了**持續的噪音**，且噪音的音量等於或大於說話聲音的音量，則必須將 <CNOISE/> 標籤放在轉錄的開頭。

例子：

- 持續的人聲噪音（笑聲、歌聲、哭泣聲等）
- 持續的環境噪音（雨水聲、交通噪音等）
- 持續的電子噪音
- 持續的背景音樂
- 持續的機械聲音
- 持續且不可理解的人群噪音

備註：如果錄音中存在**持續的噪音**，但這個噪音在錄音的大部分時間中都小於說話聲音的音量，則不應該標記 **<CNOISE/>**。

唱歌

以下的規則適用於在音頻中出現**唱歌**的片段：

- **只有歌聲的音頻**

當音頻裡只有歌聲的部分而已（或歌聲 + 音樂），那就使用標籤 **<UNKNOWN/>**。

如果整個音頻都是由歌唱，或者不可理解、被截斷或外語的說話聲音所組成，則可以使用 **<UNKNOWN/>** 標籤。

- **背景歌唱**

音頻裡出現說話聲音 + 背景歌唱的話，那就只轉錄說話聲音 + 使用 **<CNOISE/>** 標籤來標記歌唱部分。

備註：**背景歌唱**一定要用 **<CNOISE/>** 標籤。

- **間歇性歌聲**

當說話者從說話轉成唱歌（反之亦然），那就只轉錄說話聲音 + 使用 **<CNOISE/>** 標籤來標記歌唱部分。

備註：**間歇性歌聲**一定要用 **<UNKNOWN/>** 標籤。至於說話部分，則照常轉錄成文字。

<SN/>

當音頻裡出現一段短暫的噪音，並且那段噪音的音量大於說話的音量，或是和說話的音量差不多，那就必須使用 **<SN/>** 標籤。

例子：

- 人類發出的聲音（咳嗽、打噴嚏、大聲呼吸等等）
- 電子設備發出的聲音（響亮的炸裂聲、爆震聲等等）
- 機器發出的聲音（門鈴、車鳴笛等等）
- 音效

單獨出現的突發噪音

如果音頻中有**單獨出現的突發噪音**，那麼在轉錄時，應該單獨標記這個噪音，並於前後加上**空格**。

例子：

- 他即將要按門鈴 **<SN/>** 你有聽到嗎
- **<SN/>** 那裡有太多 **<SN/>**

重疊性突發噪音

如果突發噪音和單字重疊的話，那麼就應該在單字後面使用標籤，但該標籤應緊跟在重疊單字的後面，並在之後加上半角空格。

例子：

- 我一定會 <SN/> 在明天打電話給你

備註：如果音頻裡含有突發噪音，但是噪音的音量遠遠低於說話的音量，那麼就不應該標記該噪音。

備註：感嘆語和驚嘆詞，例如「嗯嗯、嗯哼、哎喲」等都不屬於噪音，因此必須要轉錄成文字。

備註：<CNOISE/> 和 <SN/> 兩個標籤都可以在同一個句子裡出現，前提是這兩個標籤在句子裡確實分別出現且必須被加上。

多個說話者和說話聲音重疊的標籤

有些音頻裡會出現兩個或兩個以上的說話者與彼此或者錄製工具進行互動和對話。

在這些情況下，無論有多少個說話者出現，轉錄員都必須依照說話的次序，將他們的談話和說話聲音逐一轉錄在同一個句子裡。

例子：

- 第一個說話者：字 1 字 2 字 3 字 4 字 5
- 第二個說話者： 字 1 字 2
- 第三個說話者： 字 1 字 2 字 3

轉錄範例：字 1 字 2 字 3 字 1 字 2 字 4 字 5 字 1 字 2 字 3

背景說話聲音

背景說話聲音是可聽見且可理解，但不是音頻裡最主要的說話聲音。

所有可以在背景中清楚辨識的話語必須依照以上規則來轉錄，包括其他說話者在同一個場景的背景裡所說的話語，以及在背景中聽到的收音機和電視裡傳出來的聲音等等。

例子：

- 第一個說話者： 字 1 字 2
- 第二個說話者： 字 1 字 2 字 3
- 背景說話聲音： 字 1 字 2 字 3 字 4

轉錄範例：字 1 字 2 字 1 字 2 字 1 字 2 字 3 字 3 字 4

針對以上所列出的狀況，如果背景聲音是無法辨識和環境造成的噪音，則必須用

<CNOISE/> 標籤。

<OVERLAP> <OVERLAP/>

使用 <OVERLAP> <OVERLAP/> 標籤來標記同時間出現兩個或兩個以上的說話者的音頻片段，包括以上所提到的可清楚辨識的背景說話聲音。

此標籤也需要用在互相重疊的外語說話聲音之間。

與其他標籤不同，這一對標籤必須用在重疊的說話聲音之間，請看以下的例子：

獨立的說話聲音 <OVERLAP> 重疊的說話聲音 <OVERLAP/> 獨立的說話聲音

對於重疊的說話聲音，需要轉錄的是最大聲、最清晰且最明顯的說話聲音。請忽略非主要的說話聲音。

如果同時出現幾個說話者，而任何一方都沒有顯得特別主要，又或者各方音量都沒有較明顯或較大聲的話，那麼請依照當下的音頻狀況，用自己的判斷力來做決定。

例子：

在這個例子中，第一個說話者的說話聲音是**最為明顯**的：

重疊

- 第一個說話者：字 1 字 2 字 3 字 4 字 5
- 第二個說話者：字 1 字 2 字 3 字 4 字 5

轉錄範例：字 1 字 2 字 3 <OVERLAP/> 字 4 字 5 <OVERLAP/> 字 3 字 4 字 5

<UNKNOWN/>、<FILL/> 和 <SN/> 標籤也可以用於 <OVERLAP/> <OVERLAP/> 標籤中，所以必須在必要時加入它們。

例子：

在這個例子中，第二個說話者的說話聲音是**最為明顯**的：

- 第一個說話者：字 1 字 2 字 3 字 4 字 5 字 6 字 7
- 第二個說話者：字 1 字 2 <FILL/> 字 3 <UNKNOWN/> 字 4

轉錄範例：字 1 <OVERLAP/> 字 1 字 2 <FILL/> 字 3 <UNKNOWN/> <OVERLAP/> 字 4

<UNKNOWN/> 和 <FILL/> 標籤被視作說話聲音。

因此，即使發生重疊的是 <UNKNOWN/> 或 <FILL/>，也必須按照規則使用 <OVERLAP/> <OVERLAP/> 標籤。

例子：

在這個例子中，第二個說話者的說話聲音是**最為明顯**的：

- 第一個說話者：字 1 字 2 字 3 <UNKNOWN/> 字 7 字 8 字 9
- 第二個說話者：<UNKNOWN/>

轉錄範例：字 1 字 2 字 3 <OVERLAP/> <UNKNOWN/> <OVERLAP/> 字 7 字 8 字 9

備註：只有在出現重疊的說話聲音情況下，才能使用 <OVERLAP/> <OVERLAP/>。如果說話者被噪音打擾，那麼在這種情況下應該使用的是噪音標籤，而不是重疊標籤。

備註：必須在標籤和說話聲音之間空格。

備註：一字不漏並且完整地轉錄一段重疊說話聲音是轉錄員常犯的錯誤。根據以上所解釋的狀況和例子，請**只**轉錄音量最明顯的說話者所說的話。

備註：**不可誤用** <OVERLAP/> <OVERLAP/> 標籤來象征「改變說話者」的情況。

其他資源

辭典

語言	辭典名稱	辭典連結
zh-TW	教育部重編國語辭典修訂本	http://dict.revised.moe.edu.tw/cbdic/
en-US	Lexico	https://www.lexico.com
en-GB	Lexico	https://www.lexico.com
en-US/GB	Merriam	https://www.merriam-webster.com/