

《臺灣客家語常用詞辭典》系統程式的維護

駱嘉鵬・李良一

清雲科技大學講師・中央大學資工系博士後研究員

新竹教育大學台語所博士生

教育部《臺灣客家語常用詞辭典》自 2008 年 5 月 15 日公告試用後，各界意見反應熱烈。為改善系統效能，自 2009 年 7 月起，委託清雲科技大學執行本辭典維護計畫。筆者駱嘉鵬為本計畫案主持人，李良一為專案程式設計師。

本維護案的主要工作內容，在於系統程式的修正。修正的重點則在改進舊系統以下缺點：邏輯錯誤、資訊不明、格式不一、限制過嚴與流程不順。至於資料內容的修正，主要由編輯小組成員負責；唯筆者亦對全典詞目音讀進行音字媒合作業，整理其中臺灣客語五大腔調音讀，製作各式音韻配合表與同音字表，以便小組成員找出問題字音，進而加以修正。

本文擬分享本專案部分工作經驗，說明其設計理念與程式修正要點，藉以徵詢學界高見。並期以此拋磚引玉，切磋琢磨，以做為爾後相關作業之參考。

關鍵字：辭典學、客語辭典、客語注音、客語用字、一致性、資料庫管理

1. 前言

教育部《臺灣客家語常用詞辭典》原由「財團法人中文資訊推廣基金會」負責設計¹，自 2008 年 5 月公告試用後，各界反應，相當熱烈。對於官方版客語網路詞典的問世，語文學界都報以熱烈支持；但在實際使用之後，對於軟體的正確性與親和力，又不免多所質疑。

為配合 2009 年底客語字音字形競賽，須有一套較為精確的標準；客語推薦用字甫研議通過，辭典資料必須配合修正；並參酌使用者意見，修正本辭典程式及內容的錯誤，以提升其使用效能，乃由教育部國語會委託清雲科技大學執行本辭典維護計畫。

本計畫名為維護，但在審視原程式之後，筆者認為除網頁畫面可沿用舊系統略加修正外，對於流程的控制與程式碼的撰寫，實不如另起爐竈，重新設計。

本辭典的修正，可分為兩個層面：一是系統程式，一是資料內容。本維護計畫主要負責程式的修正，對於資料內容的校正，也盡可能彙整相關資訊，提供編輯小組做為修正的參考。

關於本系統的修正意見，各方反應不一。除顯屬錯誤者，可由筆者或編輯小組逕行更正外，若屬仁智之見，則須經由編輯委員會議決議，方可進行修改。本文僅就本次修正所及之處加以說明。

2. 資料問題與解決方式

筆者執行本辭典維護計畫前，資料問題主要可歸為三種類型。其中用字的修改，由編輯小組先行校正；而因拼音改版或人工手誤所造成的資料問題，則由筆者進行全典詞目音讀的音字媒合作業²，藉以發掘問題字音，提供編輯小組做為修正的參考。

2.1 問題類型

2.1.1 部分用字與客語推薦用字不同

本辭典於 2008 年 5 月 16 日即公告試用，而《臺灣客家語書寫推薦用字》則至 2009 年 6 月始研議完成，後者對於前者略有修正。為使部訂標準一致，自應配合後者回校辭典用字。本項作業先由教育部國語會客語辭典編輯小組成員就用字改動部分進行修正；完成之後，再由筆者協助整理出各筆例句不含詞目之異常紀錄，提供編輯小組再次校正。

2.1.2 拼音方案修正，入聲韻尾未徹底轉換

本辭典原採用教育部於 2003 年 2 月 27 日公告之「臺灣客家語拼音方案」注音，自 2008 年 10 月 6 日起，配合本拼音方案的修訂而修正。其中主要差別，在於塞音韻尾由 p,t,k 改為 b,d,g。筆者在接手維護時，雖大部分皆已改定，但仍有多處疏漏。

2.1.3 人工手誤

本辭典涵蓋五種臺灣客語主要腔調，內容龐雜，可想而知。雖經編輯小組用心撰稿，編輯委員詳加審校，資料疏失，仍在所難免。檢視其錯誤音讀，又可歸為下述幾類：

¹ 以下所謂「舊系統」，即指該基金會所設計之系統程式；「新系統」則為筆者維護修正後之系統，亦即目前教育部網站提供服務之系統。

² 具體做法請參見駱嘉鵬，2008。

- (1) 音素缺漏，如「覓：mi55」³誤作「mi」⁴。
- (2) 音素誤增，如海陸音「合意：gag5 rhi11」誤作「gag55 rhi11」⁵。
- (3) 音序倒置，如「出口傷人：cud2 kieu31 song24 ngin11」誤作「cud2 kieu31 song24 nign11」⁶。
- (4) 空白失誤，如「毆□⁷死：eu31 bi53 xi31」誤作「eu11 bi5 3xi31」⁸。
- (5) 其他手誤，如「失漏天機：siid2 leu55 tien24 gi24」誤作「siid2 1leu55 tien24 gi24」⁹。
- (6) 綜合錯誤，如「矮人多心事：ai31 ngin11 do24 xim24 sii55」誤作「ai31 ngin11 do24 xim5 3sii55」¹⁰。

2.2 詞目音字媒合

2.2.1 前置處理

本辭典例句皆未注音，詞目音字媒合作業相對比較單純。但在媒合之前，仍有幾點必須特別處理：

- (1) 音字媒合的基本原則是依序將詞目每一個字配上音讀的一個音節。依照編輯體例，音讀欄以空白區隔音節，可做為媒合的判別基礎。但在媒合之前發現，其中仍有許多疏漏，如兩音節間遺漏空白、兩調值數字間誤加空白、包含連續空白等。此類格式錯誤，必須先行修正，以免造成音節誤判，進而導致媒合錯誤。修正的程序，是先將音讀欄所有空白清除，再將數字之後緊接英文字母者其間插入空白，如此即可將上述各類空白錯誤狀況排除。
- (2) 音讀欄包括五腔注音，須加標記處理，使其既能統合綜觀，亦能分別檢視¹¹。
- (3) 部分詞目海陸無「仔」尾，如：「梅仔樹：moi55 shu33」，其音節數自然少於詞目字數。如遇此狀況，在音字媒合前，應先將「仔」字去除，以免造成媒合錯誤。
- (4) 歇後語詞條包含破折號，在詞目中者，應先將其清除；在音讀欄者，應替換為空白。例如：「刀切薤菜——兩頭空：do24 qied2 vung55 coi55——liong31 teu11 kung24」，須先調整為：「刀切薤菜兩頭空：do24 qied2 vung55 coi55 liong31 teu11 kung24」。
- (5) 多數諺語詞條包括標點符號，在詞目中者，應先將其清除；在音讀欄者，應替換為空白。例如：「鴨嫲多，懶生卵；心白多，懶洗碗：ab2 ma11 do24，nan24 sang24 lon31；xim24 kiu24 do24，nan24 se31 von31」，須先調整為：「鴨嫲多懶生卵心

³ 本節 (2.1.3) 及 2.2.1 節注音，除另加註記者外，以四縣音為例。

⁴ 漏標調值。

⁵ 短調誤作長調。

⁶ 「nign11」，ig 倒置。

⁷ □，為四縣「分佢：bun24 gi11」之合音。

⁸ 「bi5 3」，5 3 間空白應移至其後。

⁹ 「1leu55」，音節前誤增 1 字。

¹⁰ 「xim5 3sii55」，53 之間誤增空白，3 後遺漏空白，且將四縣腔 24 調誤作海陸腔 53 調。

¹¹ 詳參下節各項字音表格式。

臼多懶洗碗：ab2 ma11 do24 nan24 sang24 lon31 xim24 kiu24 do24 nan24 se31 von31」。

- (6) 諧音雙關語包含括號與雙關字眼，應先將其清除，或重複該字眼注音。例如：「老公撥扇——淒涼（妻涼）：lo31 gung24 pad2 san55 ——qi24 liong11」，須先調整為：「老公撥扇淒涼：lo31 gung24 pad2 san55 qi24 liong11」或「老公撥扇淒涼妻涼：lo31 gung24 pad2 san55 qi24 liong11 qi24 liong11」。¹²
- (7) 少數詞條中，部分字詞加註又音，亦應先行清除，或重複該字詞。例如：詔安音「明白：men53 pa55（men53 ped43）」，須先調整為：「明白：men53 pa55」或「明白明白：men53 pa55 men53 ped43」。

2.2.2 媒合結果

詞目音字媒合結果，共得出 15 萬餘筆¹³五腔詞目字音紀錄。再按照音節內容，分析其聲韻調，製成下列各式字音表：

- (1) 字音分析表：依序將各詞條漢字與各腔調各音節媒合的初步結果，進一步分析其聲韻調而製成。格式如下：

腔調	詞目序號	字序	字	音讀	聲母	韻母	調值	字音 ¹⁴
四縣	1	1	阿	a24		a	24	阿 a24
四縣	2	1	阿	a24		a	24	阿 a24
四縣	2	2	爸	ba24	b	a	24	爸 ba24
四縣	3	1	阿	a24		a	24	阿 a24
四縣	3	2	伯	bag2	b	ag	2	伯 bag2

- (2) 字音彙總表：將上述字音分析表依照腔調別，將各腔調中同字同音之紀錄合為一筆，統計其使用頻率而製成。為便於篩檢問題字音，並分別根據各腔調值，判別調類。完成之後，再據以製作以下各表。本表五腔合計共約 23,000 餘筆，格式如下：

腔調	字音	字	音節	聲母	韻母	調值	韻調 ¹⁵	韻末	調類	筆數
四縣	一 id2	一	id2		id	2	id2	d	4	127
四縣	乙 ied2	乙	ied2		ied	2	ied2	d	4	1
四縣	八 bad2	八	bad2	b	ad	2	ad2	d	4	33
四縣	卜 bug2	卜	bug2	b	ug	2	ug2	g	4	5
四縣	丁 den24	丁	den24	d	en	24	en24	n	1	11

- (3) 聲韻調配合表：根據字音彙總表製作的音節統計表。格式如下：

腔調	(全部)
----	------

¹² 此類諧音雙關語，在音字媒合完成後，辭典體例已改為：在詞目欄只列寓意，僅在釋義欄兼註本義。

¹³ 此為五腔共同詞目部分的字音總筆數，大平安三腔特殊詞目部分的字音紀錄則近萬筆，其音字媒合仿照本文所述方式整理，不另做說明。因媒合與修正作業反覆執行，故每次媒合所得紀錄筆數略有出入。

¹⁴ 本欄係將本表之「字」與「音讀」二欄合併而成。當根據以下各配合表察覺問題字音時，搜尋本欄即可快速得知該問題字音之出處（即如腔調別、詞目序號、字序所示），而不必篩選「字」與「音讀」兩欄。

¹⁵ 本欄係將本表之「韻母」與「調值」二欄合併而成。目的為便於製作下述聲韻調配合表，以及發現問題音節時快速搜尋該類紀錄。

加總 的筆數	聲母																									
韻調	b	bb	c	ch	d	f	g	h	j	k	l	m	n	ng	p	q	rh	s	sh	t	v	x	z	zh	(空白)	總計
a11	64	4	109	110	4	237	311	152		2	12	179	16	45	59			112	17	6	7		35	21	82	1584
a113	9		55			11		20		1	2	63	16	37	21		23	17	16		3					294
a24	92		93	24	180	151	191	182		26	16	75	5	23	11		6	45	24	11	7		21	22	64	1269
a31	50	5	20	16	473	19	105	4		11	16	47	2	23	11		2	13	12		4		35	23	20	911
a33	23		19	55	5	102	137	120		12	5	48	5	8	6		83	35	22	6	5		4	5	52	757
a35						1	1							1												3
a53	91	3	88	80	165	190	231	173		9	18	121	18	50	67		57	44	60	3	11		34	23	63	1599
a55	45	20	119		5	103	51	165		15	18	154	30	106	94		23	68	94	12	13		33		6	1174

本表為 Excel 樞紐分析表格式。在左上角「(全部)」字眼處可選擇瀏覽對象為四海大平安任一腔調或五腔合計，如本例「(全部)」，即表示各統計數據為五腔合計總筆數。在任一統計數字儲存格中按兩下，即可以上述字音彙總表格式另開新表顯示該數字所代表之原始資訊。例如「bb」行「a11」列交錯的數字 4，表示「bba¹¹」音共有 4 筆，雙按該數字即可另開下列資料表：

腔調	字音	字	音節	聲母	韻母	調值	韻調	韻末	調類	筆數
詔安	蛙 bba11	蛙	bba11	bb	a	11	a11	a	1	1
詔安	榎 bba11	榎	bba11	bb	a	11	a11	a	1	3

如此則當在配合表中發現錯誤或異常數據時，即可根據「字音」欄查詢字音彙總表，進而查閱原始資訊，進行修正。以下各配合表皆仿此方式設計。

(4) 聲韻配合表：忽略聲調的音節統計表，便於查閱異常聲韻組合音節，例如：j,q,x 與洪音相配等問題。格式如下：

腔調	(全部)
----	------

加總 的筆數	聲母																									
韻母	b	bb	c	ch	d	f	g	h	j	k	l	m	n	ng	p	q	rh	s	sh	t	v	x	z	zh	(空白)	總計
a	374	32	503	285	832	814	1027	816		76	87	687	92	293	269		194	334	245	38	50		162	94	287	7591
ab			83		230	156	147	194		35	73		40	10			26	30	35	75			24	20	184	1362
ad	222	19	52	33	27	323	8	20		24	92	85	12	38	98		17	124	31	30	69		51	26	8	1409
ag	296		59	48	24	8	123	110		60	58	109	16		160		3	83	124	50	45		76	60	24	1536
ai	284	14	550		225	129	471	100		74	130	110	162	58	165		2	108		1124	43		266		85	4100
ainn										2					1											3
am			128		202	146	344	187		78	101		172	55			56	675	52	250			64	27	293	2830
an	646	89	114	12	185	464	38	253		15	126	304	265	60	320		196	713	66	206	379		95	88	42	4676

ang	62	19	44	66	143	3	200	408	38	124	211	28	110	157	86	493	299	50	90	107	114	32	2884
ann								2														1	3

(5) 韻調配合表：忽略聲母的音節統計表，便於查閱問題調值，或韻調配合異常之音節。格式如下：

腔調	(全部)
----	------

加總 的筆數	調值													
韻母	2	5	11	21	24	31	33	35	43	53	54	55	113	總計
a			1584		1269	911	757	3		1599		1174	294	7591
ab	477	384		183	127				80		111			1362
ad	495	435		157	139				97		86			1409
ag	649	536		231							120			1536
ai			495		561	508	441	1		949		1012	133	4100
ainn			2			1								3
am			622		409	398	278			632		389	102	2830
an			892		636	560	409			1045		945	189	4676
ang			716		386	151	249			584		577	221	2884
ann			1		2									3
au			595		306	395	177			712		350	36	2571

(6) 字聲配合表：各字聲母分析表，藉以發掘誤打聲母或誤用漢字，例如「乙」字讀「bb」聲母等問題。格式如下：

腔調	(全部)
----	------

加總 的筆數	聲母																										
字	b	bb	c	ch	d	f	g	h	j	k	l	m	n	ng	p	q	rh	s	sh	t	v	x	z	zh	(空白)	總計	
一																	459									127	586
乙		1															3									1	5
丁					55																						55
七			95														26										121
乜												8															8
九							41																				41
了											137															3	140
二														72													72
人												8		1617													1625
入														198													198

(7) 字調配合表：各字聲母分析表，藉以發掘誤打聲調或漢字，例如「乜」字讀第

2 聲等問題。格式如下：

腔調	(全部)
----	------

加總 的筆數	調類								
字	1	2	3	4	5	7	8	(空白)	總計
一				586					586
乙				5					5
丁	55								55
七				121					121
乜	3	1	4						8
九		41							41
了	24	111			5				140
二			57			15			72
人	6	1			1618				1625
入				1			197		198

- (8) 韻末配調表：各音節韻末與調類配合關係表，藉以發掘誤打聲調或韻母，例如「er」音讀第 1、第 7 聲等問題。格式如下：

腔調	(全部)
----	------

加總 的筆數	調類								
韻末	1	2	3	4	5	7	8	(空白)	總計
a	3189	1931	1657	255	1695	161			8888
ann	1		1	1					3
b				1680			1509		3189
d				7483			4005		11488
e	386	1244	901		283	11			2825
er	1	51			452	1			505
g				4925			2725		7650
i	4830	7122	7128	5	4581	1039		2	24707
ii	425	840	794		203	171			2433
inn	3	1	23						27
m	3102	1340	847	65	2323	100			7777
n	9031	5873	5323	1	9808	661			30697
ng	7806	3880	3196		6205	503			21590
o	2026	3940	1840	366	2641	103	2		10918
onn		1							1

u	5219	6726	4811	346	4916	570	3		22591
oo	15	11	80	1	1				108
總計	36034	32960	26601	15128	33108	3320	8244	2	155397

(9) 字音歸納表：依照聲韻調類，對字音彙總表重新歸併的結果，格式如下：

腔調	(全部)
----	------

加總 的筆數		
字	音讀	合計
一	id4	127
	rhid4	459
乙	bbied4	1
	ied4	1
	rhad4	2
	rhed4	1
丁	den1	55
七	cid4	95
	qid4	26
乜	ma1	1
	me1	2
	me2	1
	me3	4
九	giu2	41
了	a1	3
	le1	14
	le2	18
	le5	1
	leu2	15
	leu5	1
	liau2	75
	liau5	3
	lio1	7
	lio2	3
二	ngi3	57
	ngi7	15
人	min1	5
	min5	3
	ngin1	1

	ngin2	1
	ngin5	1615

(10) 同音字表¹⁶：根據字音分析表製作的同音字表。共分為六份表單，總表五腔混合，分表五腔分列，藉以發掘各腔音韻搭配異常狀況，如 j,q,x 配洪音，或四縣腔之舌尖面音、四縣以外之舌面音等。以四縣腔為例，格式如下：

韻母	調	b	p	m	f	v	bb	d	t	n	l	g	k	ng	h	j	q	x	z	c	s	zh	ch	sh	rh	零
a	11	背揸	爬划 舊耙	嗎麻 蝦蘇	華划	華		林	喇喇				卡	牙芽 倪崖 衙	哈蝦 霞蝦				查茶 捺	蛇儕						
a	24	巴爸 叭芭	芭芭 扒把	馬碼 瑪媽	花	挖	蛙	他	拿	拉		家加 佳嘉 迦伽	嚟 柳	嚟 吾	下	哈	吓		抓遮 渣植	車差 叉奢 除	社抓 沙砂 奢紗	紗袋				阿 鴉 腔
a	31	把叭	把	麼		假		打		裸		假	嚟	卡	雅	瓦			者	扯	舍捨	撒				亞 腔
a	55	巴霸 埔	怕罷 帕	罵	話化	話	哇	喀	噶	那	蚌	價假 架駕 嫁嫁	較	兒	貳	下	較	夏	炸詐 榨	扯岔 杈	社舍 續射 赦					啊
ab	2					法		答貼 搭塔	搭塔 搭	落	落	合甲 治治 鵠鵠	開	合	治	治			札紫 摺	摺	插	坡	眨			壓 鴨 押
ab	5							落	踏	納	拉	噶		噶	噶	噶			合	盒	狹	匣				
ad	2	八撥 鉢	撥 龜	抹	發 闊	幹		宜		本	烈	結		刻	嚟	嚟	瞎		哲折 扎紫	察掣	設殺	薩煞				關
ad	5	拔撥	拔 鉞	末沫 萊	活 罰	滑	猾	值	達	列	拉			軋	哈				雜	拶	徹	徹	舌	蝕		
ag	2	百伯 伯	欺擊	魄	麼	脈	賈	揀	挖			的逐	刈	魁	魁	魁			隻摘 窄	冊尺 赤拆	析閃					扼 扼 扼
ag	5					白	麥	麥		畫	割	琢	笛	捺	落	落			柵		石	碩				

上述各字音表的製作，在每次校對前執行，以便篩選出問題字音提供校正；校正完成後再重新製作，反覆執行，直到完校為止。

本辭典資料庫原僅標明調值、不標調類。但在篩檢問題字音時，調類標示法有其便利之處。因此，在製作字調配合表與韻末配調表時，皆以調類取代調值。例如某字在各腔的調值分別為 31,24,31,53,11，不易看出有何問題；但若改以調類分別標註第 2,2,2,2,1 聲，則可發現末者異常。又如「十」字各調值總計筆數如下：

調值	2	5	43	54
筆數	32	59	26	29

上表無法察覺有何問題。但若分別歸入各腔調類，重新統計筆數如下：

調類	4	8
筆數	1	145

如此則可明顯看出第 4 聲該筆異常，即可據此予以修正。

3. 程式問題與解決方式

3.1 邏輯錯誤

3.1.1 音讀模糊檢索結果排除精確紀錄

模糊（廣義）檢索的功能，主要是讓使用者在限制較小的條件下，查到更多的資訊；抑或在部分條件不明的狀況下，查到可能需要的資訊。例如：以精確檢索方式查詢詞目「八角」，僅能查到一筆；若改以模糊檢索方式查詢，則可查到「八角、八角樓、八角

¹⁶ 具體製作方法請參見駱嘉鵬，2006。

亭」等三筆。音讀模糊檢索，理當仿此原則設計。舊系統以精確檢索方式查詢「bad2 gog2」，可以查得「八角」；但以模糊檢索方式查詢時，則僅查到「八角樓、八角亭」兩筆。

新系統依照資料庫正規檢索邏輯設計，已排除本項錯誤。

3.1.2 部分圖形字輸出異常

系統對於 **unicode** 缺字或某些作業系統無法正常顯示之字，皆以圖形字來顯現。唯舊系統處理程序有誤，某些圖形字會變成亂碼或「？」號，而無法正常顯示。例如查詢詞目「拈甜」，點按該筆內容如下：

詞目	【拈甜】	詞性：動	各家用字表
四縣音	ngiam ²⁴ tiam ¹¹		
海陸音	ngiam ⁵³ tiam ⁵⁵		
大埔音			相關資料連結
饒平音			相關資料連結
詔安音			相關資料連結
釋義	指過年或逢喜事時，來訪的客人要拿取糖果等甜點吃的習俗：阿爸交代，過年時節有人客來，愛請佢～甜（爸爸交代，過年的時候有客人來玩，要請他們吃一些甜點）。		
近義詞			
反義詞			
文白讀			
又 音			
多音字			
對應華語			
圖片			

上例「个、寮」皆變為「？」號。新系統已修正該項錯誤，所有圖形字皆能正常顯示。如上例，瀏覽畫面修正如下：

詞目	【拈甜】	詞性：動	各家用字表
四縣音	ngiam ²⁴ tiam ¹¹		
海陸音	ngiam ⁵³ tiam ⁵⁵		
大埔音	ngiam ³³ tiam ¹¹³		
饒平音	ngiam ¹¹ tam ⁵⁵		
詔安音	ngiam ¹¹ tem ⁵³		
釋義	指過年或逢喜事時，來訪的客人要拿取糖果等甜點吃的習俗：阿爸交代，過年个時節有人客來寮，愛請佢拈甜。（爸爸交代，過年的時候有客人來玩，要請他們吃一些甜點。）		
近義詞			
反義詞			
文白讀			
又 音			

多音字	
對應華語	
圖片	

3.1.3 「多音字」與「又音」超連結，偶有疏漏

對於一字多音或一詞形多音讀之詞條，系統分列多筆詞目，並在各筆瀏覽視窗加註其他同詞形各筆音讀資訊。若為單字詞目，則加註在「多音字」欄；多字詞則加註在「又音」欄；並在加註欄內各音讀訊息上分別建立與該音讀紀錄之連結，以便使用者即時查詢。舊系統將此連結訊息另立欄位，直接存入剔除該筆以外的所有又音，以人工建立，難免疏漏；尤其在注音、用字有所調整時，更難免顧此失彼。且其超連結訊息，亦多錯誤，以致無法正常連結。

例如「著」字有六種音讀，以四縣「cog5」音為例，該筆紀錄瀏覽畫面「多音字」欄加註其他五音如下：

詞目	【著】	部首:艸	筆畫:8-12
四縣音	(一) cog		
海陸音	chog`		
大埔音			相關資料連結
饒平音			相關資料連結
詔安音			相關資料連結
釋義	1. 正確的、對的：～毋～。 2. 開始：～手。 3. 火燃燒起：～火。 4. 遭受、受到：～驚。 5. 抽中、考中：～獎。 6. 必須、只好：～用。		
近義詞			
反義詞			
文白讀			
又 音			
多音字	(二) dau11 四 dau55 海〔見【著】條〕 (三) diau11 四 diau55 海〔見【著】條〕 (四) do31 四 do24 海〔見【著】條〕 (五) zog2 四 zhog5 海〔見【著】條〕 (六) zu55 四 zhul1 海〔見【著】條〕		
對應華語			
圖片			

新系統資料庫刪除「多音字」欄，畫面中該欄資料係由程式動態搜尋現有資料庫中同詞形之紀錄組合而成，其瀏覽畫面如下：

詞目	【著】	部首:艸	筆畫: 08-12
四縣音	(一) cog		
海陸音	chog`		
大埔音	chog`		
饒平音	chog		

詔安音	choo										
釋義	1. 正確的、對的：著毋著（對不對）？ 2. 火燃燒起：著火（著火，比喻生氣）。 3. 抽中、考中：著獎。 4. 確定：定著。 5. 遭受、受到、得到：著病。 6. 必須、只好：著用。 7. 開始：著手。										
近義詞											
反義詞											
文白讀											
又 音											
多音字	<table border="1"> <tr> <td>(二) dau^ˇ 四 dau 海</td><td>〔見【著】條〕</td></tr> <tr> <td>(三) diau^ˇ 四 diau 海</td><td>〔見【著】條〕</td></tr> <tr> <td>(四) do^ˇ 四 do^ˇ 海</td><td>〔見【著】條〕</td></tr> <tr> <td>(五) zog^ˇ 四 zhog 海</td><td>〔見【著】條〕</td></tr> <tr> <td>(六) zu 四 zhu^ˇ 海</td><td>〔見【著】條〕</td></tr> </table>	(二) dau ^ˇ 四 dau 海	〔見【著】條〕	(三) diau ^ˇ 四 diau 海	〔見【著】條〕	(四) do ^ˇ 四 do ^ˇ 海	〔見【著】條〕	(五) zog ^ˇ 四 zhog 海	〔見【著】條〕	(六) zu 四 zhu ^ˇ 海	〔見【著】條〕
(二) dau ^ˇ 四 dau 海	〔見【著】條〕										
(三) diau ^ˇ 四 diau 海	〔見【著】條〕										
(四) do ^ˇ 四 do ^ˇ 海	〔見【著】條〕										
(五) zog ^ˇ 四 zhog 海	〔見【著】條〕										
(六) zu 四 zhu ^ˇ 海	〔見【著】條〕										
對應華語	正確、著、中、定、得										
圖片											

至於「著著」一詞亦有六種音讀，舊系統雖可搜尋到該六筆紀錄，但若點選其中任何一筆，則皆出現錯誤訊息，而無法正常連結。

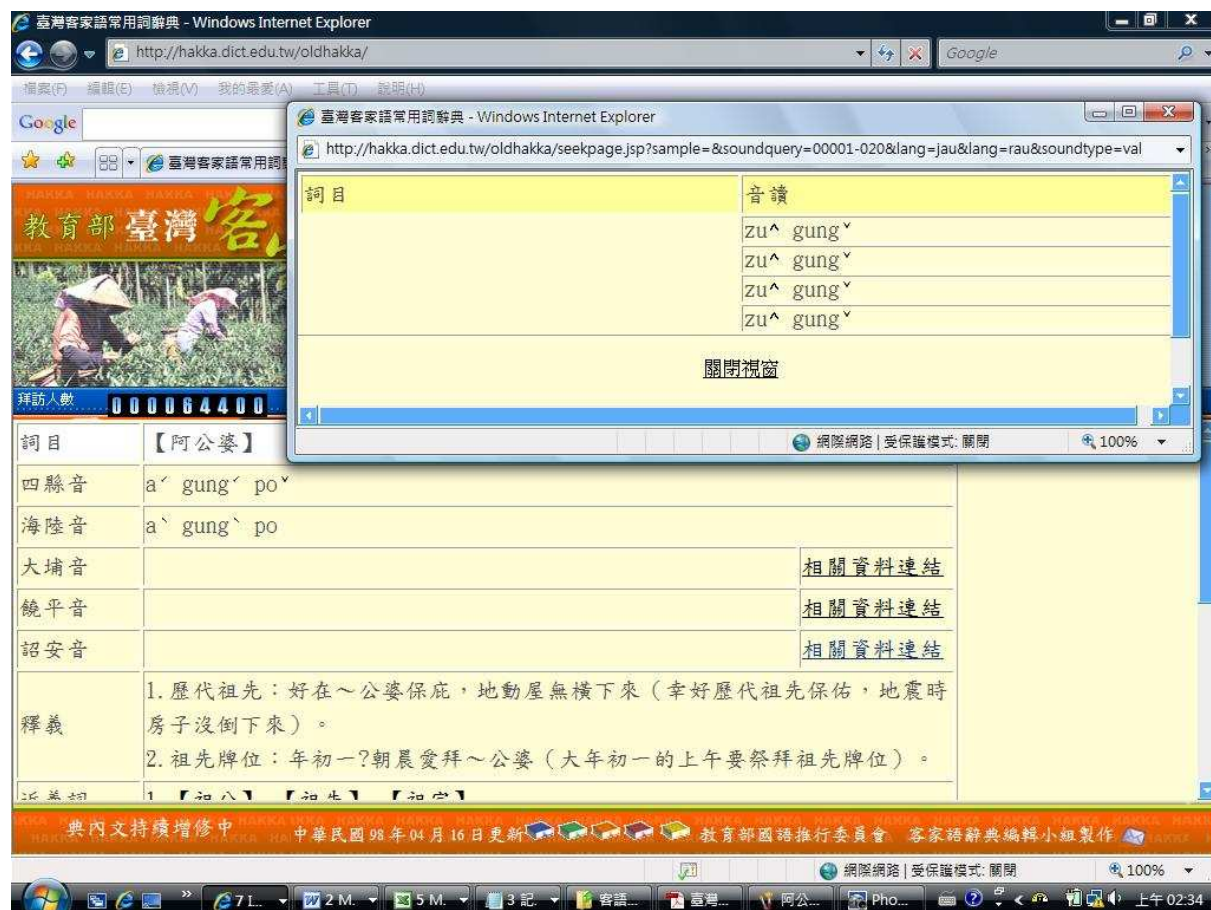
新系統則如同單字同形異音詞設計模式，差別僅在其他各音及其連結設定係以程式動態處理後填入畫面「又音」欄，並且皆能正常連結。其瀏覽畫面如下：

詞目	【著著】	詞性：形	各家用字表未收錄
四縣音	cog cog		
海陸音	chog` chog`		
大埔音	chog` chog`		
饒平音	chog chog		
詔安音	choo choo		
釋義	完全對、正確：這擺个揣令仔比賽，逐個全分佢揣 著著 ，所以〔故所〕佢拿著最大獎。（這次的猜謎比賽，每個全給她猜中，所以她拿到最大獎。）		
近義詞			
反義詞			
文白讀			
又 音	◆cog do`四 chog` do`海	〔見【著著】條〕	
	◆dau` do`四 dau do`海	〔見【著著】條〕	
	◆diau` diau`四 diau diau 海	〔見【著著】條〕	
	◆diau` do`四 diau do`海	〔見【著著】條〕	
	◆zog` do`四 zhog do`海	〔見【著著】條〕	
多音字			
對應華語	正確		

圖片	
----	--

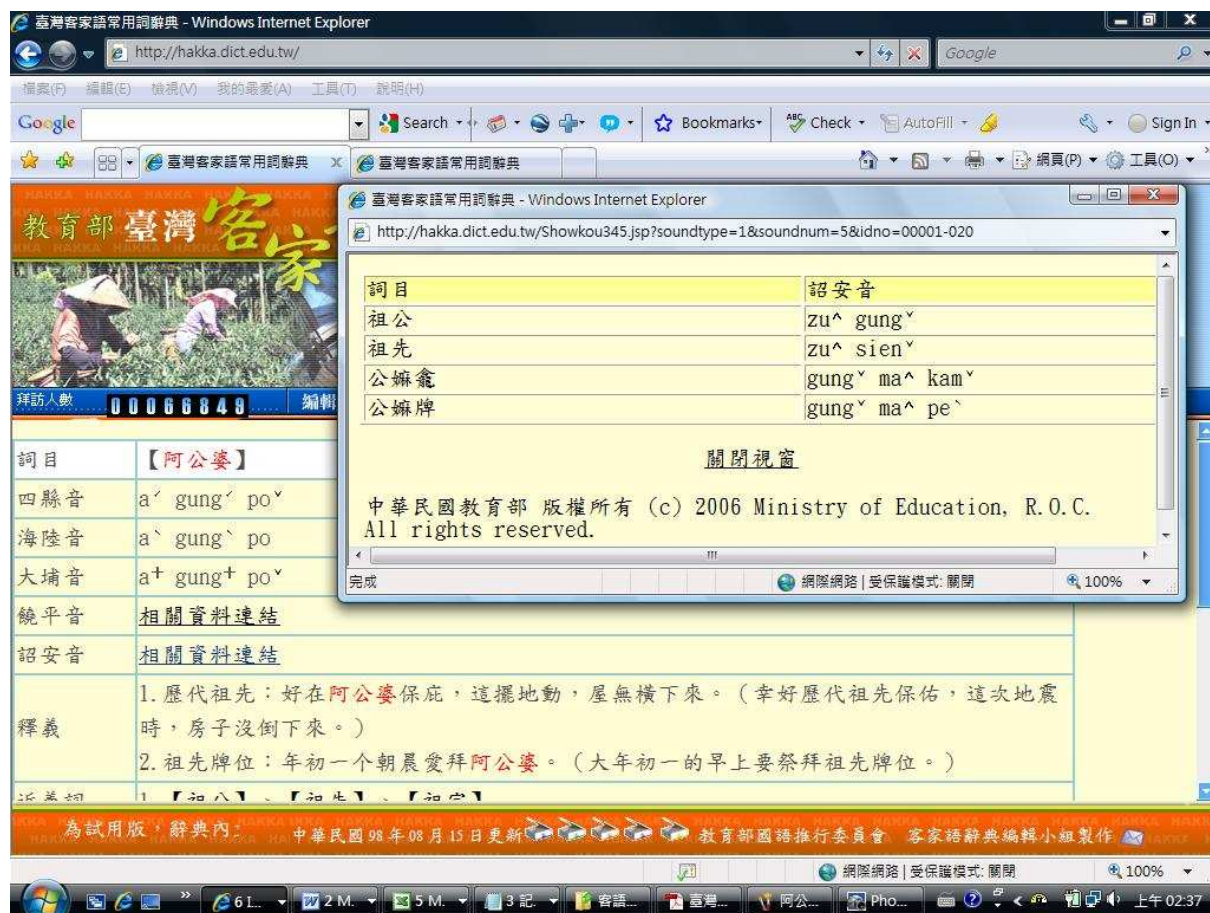
3.1.4 大平安三腔相關資料連結視窗，一對多者，音讀連結錯誤

對於大平安三腔音讀，舊系統皆以「相關資料連結」提示，待使用者點按之後，再另開視窗呈現其音讀。唯若相對於主詞目，該腔調說法超過一種時，該視窗雖分列標註音讀，但每列音讀內容皆與首列相同，顯有失誤。例如瀏覽「阿公婆」一詞時點按詔安音右側之「相關資料連結」提示後，畫面如下：



上列畫面中，「相關資料連結」視窗中之音讀，四筆皆為「zu^ gung^」。

新系統重新設計檢索流程，已排除此類錯誤。同樣瀏覽「阿公婆」一詞時點按詔安音右側之「相關資料連結」提示按鈕後，畫面修正如下：



3.2 資訊不明

3.2.1 釋義欄中省略符號，破壞資訊的完整性，造成檢索結果的疏漏

舊系統沿襲紙本辭典常例，在釋義內容中，將詞目首字以「～」號代替，以資醒目。如「鴨春」條釋義如下：

鴨蛋：～春比雞卵較大（鴨蛋比雞蛋大一些）。

如此瀏覽該筆固然醒目，但若以模糊檢索方式查詢釋義欄之「鴨春」，則因「鴨」字已被替換為「～」，而無法查得。嚴重影響釋義欄檢索的準確性。

新系統在資料庫中保留完整資訊，對於使用者輸入的檢索字詞，則以不同顏色字形標示。如此不但具有醒目效果，且不限定在詞目首字，完全根據使用者輸入的查詢字眼來標示字形色彩。如上例瀏覽畫面如下：

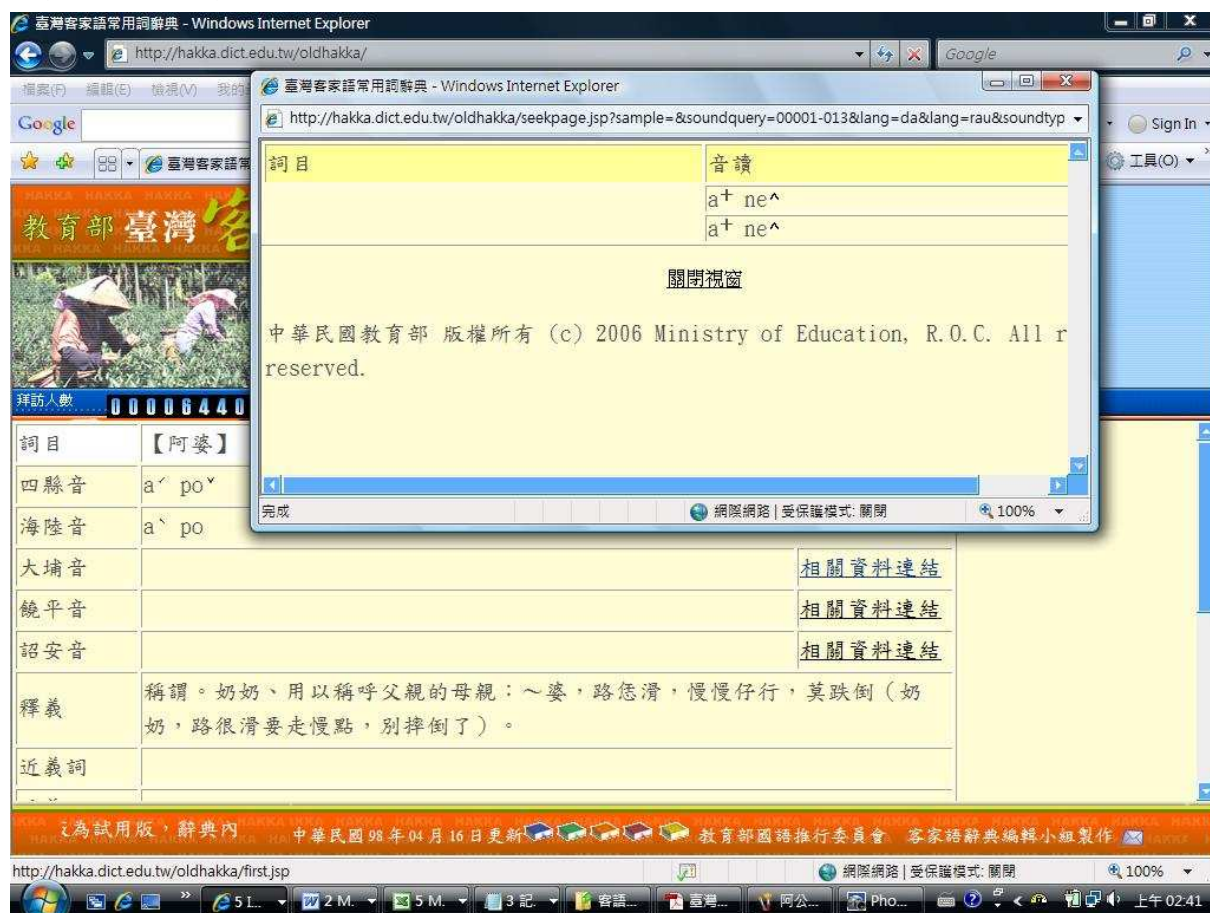
	【鴨春】	詞性：名	各家用字表
四縣音	ab` cun`		
海陸音	ab chun`		
大埔音	ab^ chun^		
饒平音	ab` chun`		
詔安音	ab` chin`		
釋義	鴨蛋：鴨春比雞卵大一息仔。（鴨蛋比雞蛋大一些。）		
近義詞	【鴨卵】		
反義詞			

文白讀	
又 音	
多音字	
對應華語	鴨蛋
圖片	

且因資料庫內容完整，並未將詞目首字換成「～」號，故以「鴨春」查詢釋義，自然也能查到該筆紀錄。

3.2.2 大平安三腔相關資料連結視窗，未標示腔調別及詞目

舊系統點按「相關資料連結」紐後，新開視窗僅註明音讀，而未標示該音讀所屬腔調，以及該音讀對應詞目。如在瀏覽「阿婆」條時點按大埔音右方之「相關資料連結」紐後，畫面如下：



The screenshot shows the 'Hakka Dictionary' website in Internet Explorer. The main page displays the search result for '阿婆' (Apo) with various pronunciations (四縣音, 海陸音, 大埔音, 饒平音, 詔安音) and a definition. A pop-up window titled '相關資料連結' (Related Links) is open, showing a table with two columns: '詞目' (Word) and '音讀' (Pronunciation). The table lists two entries with the pronunciation 'a+ ne^'. The pop-up window also includes a '關閉視窗' (Close Window) button and a copyright notice for the Ministry of Education, R.O.C. (2006).

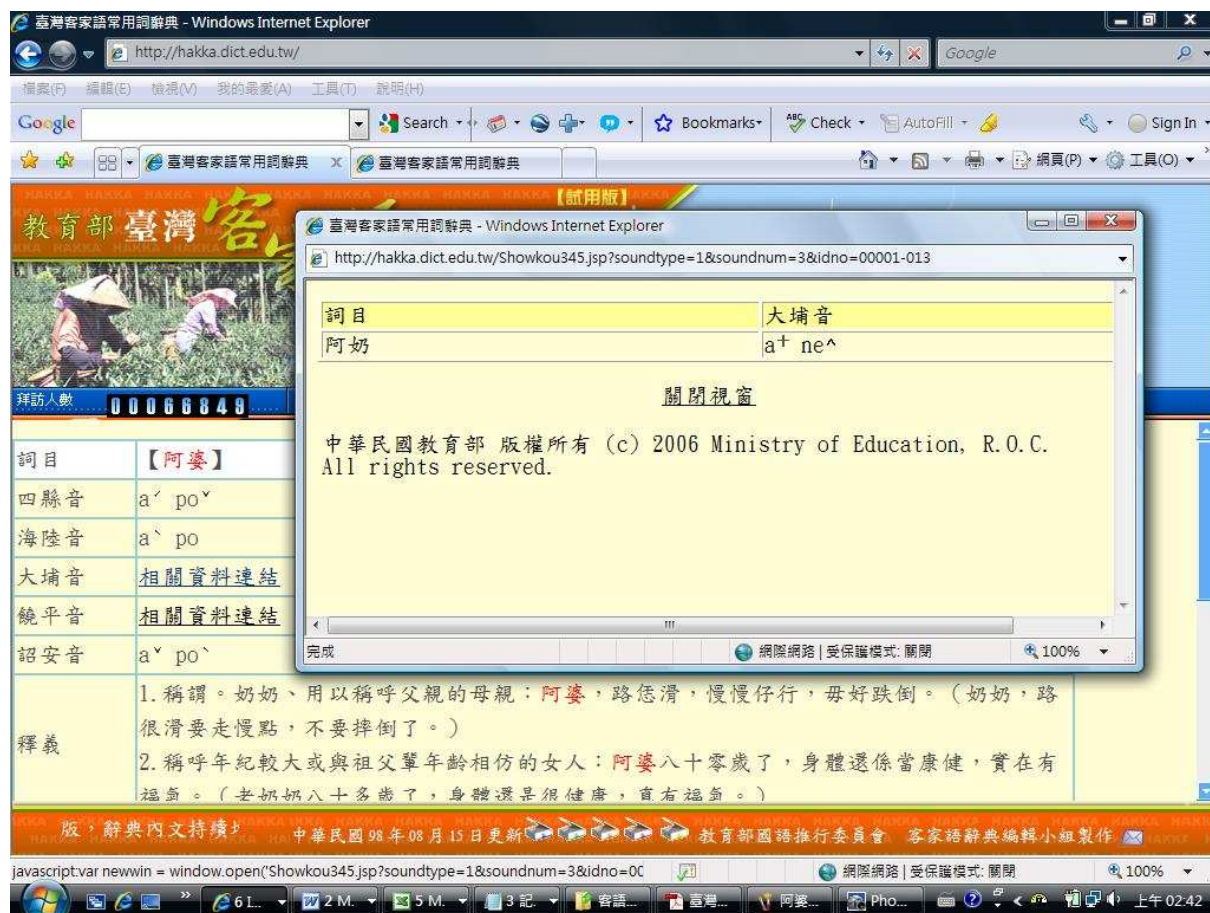
詞目	音讀
	a+ ne^
	a+ ne^

關閉視窗

中華民國教育部 版權所有 (c) 2006 Ministry of Education, R.O.C. All rights reserved.

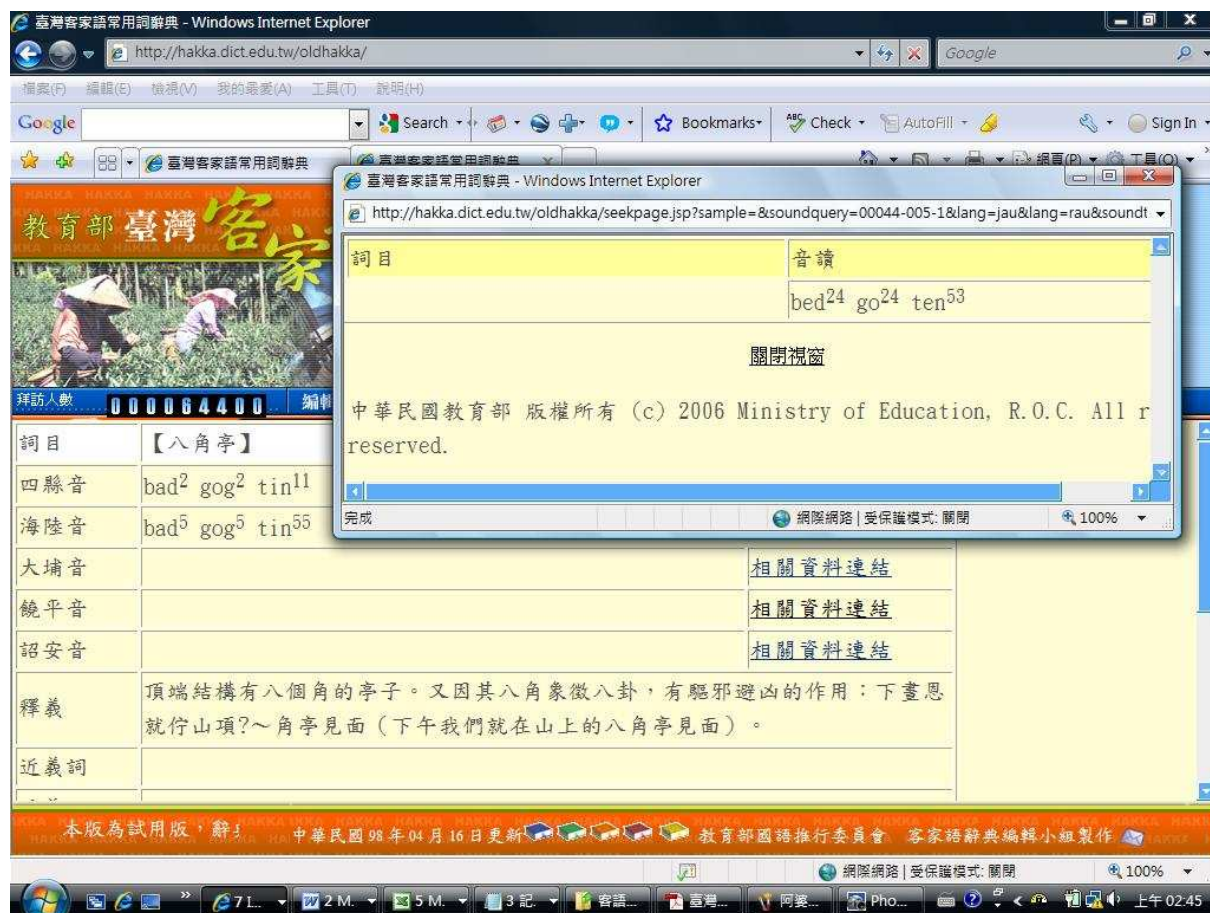
從以上畫面不能得知該「a+ ne^」音係屬大平安之何種腔調？亦不可得知該音讀對應之漢字詞目為何？

新系統修正以上缺點，在音讀欄上方標註腔調別，在各音讀左側標註對應之詞目。如上例改以下列方式呈現：



3.2.3 入聲複調調值未加底線

依趙元任先生五度制標示法，若短調為平調，調值係以單一數字標示；若非平調，則在數字下方標註底線，以與長調區別。舊系統入聲複調調值皆未加上底線，如查詢「八角亭」後點按詔安音右側之「相關資料連結」畫面如下：



新系統修正上述錯誤，如上例修正為：

詞目	【八角亭】	詞性：名	各家用字表未收錄
四縣音	bad² gog² tin¹¹		
海陸音	bad⁵ gog⁵ tin⁵⁵		
大埔音	bad²¹ gog²¹ tin¹¹³		
饒平音	bad² gog² tin⁵⁵		
詔安音	bed²⁴ go²⁴ ten⁵³		
釋義	頂端結構有八個角的亭子；其八角象徵八卦，有驅邪避凶的作用：下晝𪗇就 在山頂个八角亭見面。（下午我們就在山上的八角亭見面。）		
近義詞			
反義詞			
文白讀			
又 音			
多音字			
對應華語	八角亭		
圖片			

3.3 格式不一

3.3.1 大平安三腔音讀呈現方式，與四海二腔明顯區隔

舊系統查詢結果詞目列表畫面，當游標移至任一詞目時，即浮現該詞各腔音讀視

窗，但大平安三腔雖標示欄位，實際音讀則皆為空白。單筆詞目瀏覽畫面，四海二腔直接標註音讀，大平安三腔則僅以「相關資料連結」提示，使用者點按之後，方可查得音讀。例如查詢「阿公太」時，在檢索結果列表畫面，當游標移至該詞目時畫面如下：

編號	詞目
1	【阿公太】

四縣音：a^ˊ gung^ˊ tai^ˊ
 海陸音：a^ˋ gung^ˋ tai^ˋ
 大埔音
 饒平音
 詔安音

瀏覽該詞時畫面如下：

詞目	【阿公太】	各家用字表
四縣音	a ^ˊ gung ^ˊ tai ^ˊ	
海陸音	a ^ˋ gung ^ˋ tai ^ˋ	
大埔音		相關資料連結
饒平音		相關資料連結
詔安音		相關資料連結
釋義	稱謂。曾祖父，稱呼父親的祖父：～公太歲數恁大，身體還當健（曾祖父的年紀很大但身體還很健康）。	

新系統對於大平安三腔音讀，以下述原則處理：若詞目與四海二腔相同，則不論檢索結果音讀浮現視窗或單筆詞目瀏覽視窗皆與四海二腔並列，不另做連結。若任一腔調詞目與四海二腔不同，或為多對一關係，則該腔調音讀依舊系統模式處理：在檢索結果音讀浮現視窗不顯示，單筆詞目瀏覽視窗提示「相關資料連結」，使用者點按該連結按鈕時再另開視窗標示該腔調詞目及音讀。如上例在檢索結果列表畫面，當游標移至該詞目時畫面如下：

編號	詞目
1	【阿公太】

四縣音：a^ˊ gung^ˊ tai^ˊ
 海陸音：a^ˋ gung^ˋ tai^ˋ
 大埔音：a^ˊ gung^ˊ tai^ˋ
 饒平音：a^ˋ gung^ˋ tai^ˋ
 詔安音：

中華民國教育部 版權所有

點按該詞時畫面如下：

詞目	【阿公太】	詞性：	各家用字表
四縣音	a' gung' tai		
海陸音	a` gung` tai`		
大埔音	a+ gung+ tai`		
饒平音	a` gung` tai`		
詔安音	相關資料連結		
釋義	稱謂。曾祖父，稱呼父親的祖父：阿公太歲數恁大，身體還當健。（曾祖父的年紀很大但		

舊系統大平安三腔皆須再按「相關資料連結」提示，才可查看其音讀；新系統則只有詔安腔詞目不同，才需另開視窗查閱。

3.3.2 調值、調型選項功能，不夠完備

舊系統除檢索結果音讀浮現視窗外，調值皆未上標。除檢索結果音讀浮現視窗及大平安三腔相關資料連結視窗外，不能顯示調型。例如設定以「調型」顯示，再查詢「中央」一詞時，點按該詞條內容如下：

詞目	【中央】	詞性：副	各家用字表
四縣音	dung' ong'		
海陸音	dung` ong`		
大埔音			相關資料連結
饒平音			相關資料連結
詔安音			相關資料連結
釋義	居於中間或中心部位：該招牌～央寫誼係麼？字（那招牌中央寫的是什麼字）？		
近義詞	【中間】 【中心】		
反義詞			
文白讀	◆zhung11 rhong11 饒 文		
又 音	zung24 iong24 四 zhung53 rhong53 海 〔見【中央】條〕。		
多音字			
對應華語	中央		
圖片			

其中「文白讀」及「又音」部分仍標示調值，並未隨使用者選擇而改變，且其數字亦未上標。

新系統將調值、調型選項功能擴大適用到所有可能出現聲調的欄位，所有調值皆採用上標字形標示。如上例畫面修正如下：

詞目	【中央】	詞性：副	各家用字表
四縣音	dung' ong'		
海陸音	dung` ong`		
大埔音	dung+ ong+		
饒平音	dung' ong' 白		
詔安音	dung' ong'		

釋義	居於中間或中心部位：該招牌 中央 寫誤係麼个字？（那招牌 中央 寫的是什麼字？）
近義詞	【中間】、【中心】
反義詞	
文白讀	◆zhung ^v rhong ^v 饒 文
又 音	◆zung ^ˊ iong ^ˊ 四 zhung ^ˋ rhong ^ˋ 海 [見【中央】條]
多音字	
對應華語	中央
圖片	

若設定以調值顯示，則畫面如下：

詞目	【中央】	詞性：副	各家用字表
四縣音	dung ²⁴ ong ²⁴		
海陸音	dung ⁵³ ong ⁵³		
大埔音	dung ³³ ong ³³		
饒平音	dung ¹¹ ong ¹¹ 白		
詔安音	dung ¹¹ ong ¹¹		
釋義	居於中間或中心部位：該招牌 中央 寫誤係麼个字？（那招牌 中央 寫得是什麼字？）		
近義詞	【中間】、【中心】		
反義詞			
文白讀	◆zhung ¹¹ rhong ¹¹ 饒 文		
又 音	◆zung ²⁴ iong ²⁴ 四 zhung ⁵³ rhong ⁵³ 海 [見【中央】條]		
多音字			
對應華語	中央		
圖片			

3.4 限制過嚴

3.4.1 音讀模糊查詢，不得省略調值

舊系統以模糊檢索方式查詢音讀時，雖可僅輸入詞條中任一字音，但該注音必須包括完整音節，不得省略調值。如輸入「ngin」無法查得「ngin11、ngin55、qid2 ga24 ngin11.....」等任一筆紀錄，需輸入「ngin11」才可查得「ngin11、qid2 ga24 ngin11」等紀錄。由於調值標示法，使用者未必熟悉；調值的細微差異，又往往是見仁見智。如此設計，難免造成不便。

新系統模糊查詢完全依照資料庫檢索通則設計，只要檢索對象包含使用者輸入字串，即視為已經檢得。例如使用者查詢「ngin」，則不論「ngin11、ngin55、ngin113、ngin53、qid2 ga24 ngin11.....」皆可查得；若使用者查詢「ngin11」，則依然可限定檢索條件為包含「ngin11」之「ngin11、qid2 ga24 ngin11.....」等紀錄。

3.5 流程不順

3.5.1 單筆詞目瀏覽視窗，畫面捲動遲滯

舊系統在瀏覽單筆詞目內容時，畫面向下捲動，常有遲滯現象。新系統重新設計，已無此問題。

3.5.2 大平安三腔相關資料連結視窗開啟，耗時過久

舊系統在點按大平安三腔「相關資料連結」按鈕後，新視窗的開啟，往往耗時過久；有時甚至導致當機。新系統業已修正。

4. 餘論

筆者執行本辭典維護案，迄今大約兩個月，僅先針對上述重大問題進行修正。此外，仍有諸多改進空間。例如釋義欄中，「僅四縣用」、「僅海陸用」等行文用語，原僅考慮海陸二腔，如今五腔並列，多半必須加以修改。本次修正，限於人力，未及全面校正。此類資料問題，尚待編輯小組成員陸續更正。

其次，關於例句用字，本次亦未及處理。若能從釋義欄內抽離所有例句，全面彙整其用字，再與詞目用字詳加比對，應有助於再次檢驗兩者用字之一致性，並可整理出例句用字（詞）超出詞目之外者，以做為更正用字或增補詞目的參考。

再者，辭典收詞原則，似仍有討論餘地。又如用字問題，閩客語之間存在許多異於華語的同源詞，或相互影響的借詞。這類用字，若能取得一致，對於雙語交流學習，亦必具有促進之功效。唯此則有賴閩客語學者共同努力，高瞻遠矚，捐棄成見，方可順利達成。

相較於《臺灣閩南語常用詞辭典》，本辭典例句皆未注音。此固由於五腔並列有所困難，但對於不諳客語之初學者而言，無疑是一大阻礙。未來若能進而連結聲音檔，必能產生更大助益。

本系統的程式功能，亦必仍有值得商榷之處，此則謹請與會專家惠賜高見。若為系統失誤，自應力求改進；若屬仁智之見，則可提供編輯委員會討論，以做為系統改版的依據。

參考文獻

- 教育部國語推行委員會，2008a，《臺灣客家語常用詞辭典(試用版)》，
<http://hakka.dict.edu.tw/>。
- ，2008b，《臺灣閩南語常用詞辭典(試用版)》，<http://twblg.dict.edu.tw/tw/>。
- ，2009a，《臺灣客家語拼音方案》，
http://www.edu.tw/mandr/bulletin.aspx?bulletin_sn=3824&pages=0&site_content_sn=3352。
- ，2009b，《臺灣客家語書寫推薦用字》(第1批)，
<http://www.edu.tw/files/download/M0001/hakka-words.pdf>。
- 駱嘉鵬，2006，〈漢字同音字表的新做法——以臺灣閩南語字音表為例〉，廈門大學，第38屆漢藏語國際學術研討會。<http://w3.cyu.edu.tw/luo/jiapeng/d1-6.htm>。
- ，2008，〈《臺灣閩南語常用詞辭典》音字系統一致性的校正——資料庫管理模式在辭典校對作業上的用例〉，臺灣師範大學，第七屆臺灣語言及其教學國際學術研討會。<http://w3.cyu.edu.tw/luo/jiapeng/d1-15.pdf>。