



TAFT: A new version of pinyin to help foreign Mandarin learners remember tones

Marcus Taft, Beijing Language and Culture University & UNSW Sydney 

Abstract

A group of 30 native English-speaking participants were taught in a single session how to produce tones in Mandarin Chinese and memorized a set of 24 English words along with their Chinese translation presented in an alphabetic script. For half of the participants, the alphabetic script took the form of the standardly used “pinyin” where tonal information is depicted by a diacritic line above the vowel. For the other half, the script represented the tones as letters such that each syllable and its tone formed an integrated unit. This method was named “Tones as Alphabetically Formed Tokens” or “TAFT”. Using a cued-retrieval task where the Mandarin word was to be pronounced in response to its English translation, it was found that the TAFT method was significantly better than standard pinyin for retaining both the syllable and its tone. This was true when testing both immediately after learning and after a week’s delay. It is suggested that TAFT is potentially a more effective tool for mediating the teaching of Mandarin as a foreign language than is the pinyin script that is currently used.

Keywords: Pinyin, Teaching Mandarin as a Foreign Language, Learning Lexical Tones, Romanization of Chinese

APA Citation: Taft, M. (2024). TAFT: A new version of pinyin to help foreign Mandarin learners remember tones. *Second Language Research & Practice*, 5(1), 133-144. <https://hdl.handle.net/10125/69887>

Introduction

Chinese comprises a family of languages that are all tonal in nature, whereby the meaning of a word is different depending on the pitch level. Mandarin Chinese uses four different tones: Level pitch (Tone 1), rising pitch (Tone 2), dipping pitch (Tone 3), and falling pitch (Tone 4). For example, the syllable /fu/ means “fortune” when spoken with Tone 2, but “rotten” when spoken with Tone 3.

Learning tones in L2

When learning Chinese as a second language, it is a major problem to be able to perceive the correct tone of a word as well as to produce it when pronouncing the word aloud, especially when the learners’ L1 is non-tonal. Research into L2 tonal learning has mainly examined the ability of the learner to identify the correct tone (e.g., Godfroid, et al., 2017; Liu, et al., 2011; Pelzl, 2019; Showalter & Hayes-Harb, 2013) but the present study focuses on production. Here, there are actually two problems for the learner. First, there is the difficulty in being able to accurately articulate the appropriate tone (see Třísková, 2017, for discussion of how tonal production might be taught more effectively). The second problem is even more basic, though, and that is the difficulty in remembering which is the correct tone to use for any particular word. Even if an L2 speaker is able to accurately pronounce Mandarin tones, they will still have difficulty

communicating if they assign the wrong tone to the word they are trying to produce. The focus of the present research is specifically on this second problem.

Using pinyin to represent tones

When teaching Mandarin, textbooks currently use the alphabetic transcription that was developed in China as a mediating tool for teaching native speakers how to read Chinese characters. This is called ‘Hanyu Pinyin’, or ‘pinyin’ for short. In the pinyin system, the phonemic segments are represented by letters of the Roman alphabet and tones are represented by a digit (1, 2, 3, and 4) or, more typically, by a diacritic line drawn above a vowel to indicate whether it is level (i.e., –), rising (i.e., /), dipping (i.e., v), or falling (i.e., \). Thus, ‘fú’ (with the rising Tone 2) means “fortune”, while ‘fǔ’ (with the dipping Tone 3) means “rotten”.

Since pinyin is the medium through which non-native speakers typically learn the correct pronunciation of Mandarin words, it is important for the written transcription to help learners retain the correct pronunciation of the word in lexical memory. This is where a problem arises with diacritics being used to indicate tone.

In particular, when developing their vocabulary, learners not only have to remember the appropriate phonemic segments but also the tone that goes with the syllable, and these two types of information are conveyed in pinyin quite differently. That is, phonemes are represented by letters, and tones are represented by diacritics. The learner must therefore integrate these two types of information in order to remember the correct pronunciation of the word, and there are a number of reasons to expect this to be quite difficult.

First, native speakers of alphabetically scripted languages that do not use diacritics (e.g., English) are highly familiar with the use of letters to represent verbal information, but not with the use of lines to do so. When attempting to gain new knowledge, it is easier to hold familiar information taken from long-term memory in working memory than unfamiliar information (e.g., Sweller, 2022). On top of this, the involvement of the two types of visual information to be memorized, rather than one, may add to the cognitive load by splitting attention between the two different categories of information (see e.g., Sweller, et al., 1990).

There is also a neurophysiological basis for suggesting that it might be hard to integrate the letters and lines when using pinyin to learn the pronunciation of new words since the processing of written words involves different brain mechanisms to the processing of shapes that are, in themselves, non-verbal tokens. The former involves the ventral visual pathway and the latter the dorsal visual pathway (e.g., Grill-Spector & Malach, 2004), while the visual word form area that is specialized for the processing of written words, is quite a different region of the brain to where shapes are processed (e.g., Dehaene & Cohen, 2011).

So, while pinyin might provide a useful way to represent tones for people who already know the pronunciation of Mandarin words, it may not be the optimal way of representing tones for non-native learners, especially if their L1 does not use diacritics. That is, it may well be hard to retain the tonal and phonemic information as an integrated unit when they are each represented by a different category of visual information. We can therefore ask if there is a better way to visually represent the phonemic and tonal information of Mandarin such that memory for the correct pronunciation is facilitated.

Tonal spelling

What can be proposed is that Roman letters are not only used to represent phonemes in the pinyin script but are also used to represent tones. The use of the same visual category for both phonemes and tones would allow a single unit to be retained when learning new Mandarin words. Such an idea was indeed a feature of the early romanization of Chinese called Gwoyeu Romatzyh (GR), developed by Yuen Ren Chao and others in the 1920s. This writing system was officially adopted in China in 1928 and used in many teaching materials in the West, though subsequently abandoned (see Kratochvíl, 1968). In GR, tonal spelling was adopted whereby the level tone (Tone 1) was taken to be the basic form and the other tones were generated from this by changing the vowel to another vowel, adding a silent letter, or doubling a letter. However, there was considerable complexity in the way these variations were used to generate the different tones, with the particular application depending on the identity of the vowel, the existence of a consonantal coda, and the type of consonantal onset. As a result, the relationship between the spelling of the different tones was quite opaque. For example, ‘ching’, ‘chang’, and ‘chuan’ all represent words with Tone 1, while ‘chyng’, ‘charng’, and ‘chwan’ represent the same syllables with Tone 2.

Clearly, such complexity works against the rationale for using tonal spelling as a means of facilitating the learning of tones. Indeed, a study by McGinnis (1997) demonstrated the inefficacy of GR when it was compared with standard pinyin as a teaching medium. Each medium was used to teach one year of an elementary level Chinese course in the USA. At the end of the course, it was found that those taught through GR were no more accurate in their tonal production, as judged by native Chinese speakers, than those taught through standard pinyin. If anything, they were poorer.

The only other published suggestion for representing tones in Mandarin with letters is the largely ignored proposal of Anderson (1967), which he calls ‘Simplified Wade’. This is a far more systematic approach than GR. A different consonant is added to the end of the syllable to indicate each tone. Tone 1 is again taken as the basic tone and has no letters added (e.g., ‘ba’, ‘ting’), while the other three tones are indicated by an added consonant that never otherwise occurs in final position in pinyin: A final ‘v’ for Tone 2 (e.g., ‘bav’, ‘tingv’), a final ‘x’ for Tone 3 (e.g., ‘bax’, ‘tingx’), and a final ‘z’ for Tone 4 (e.g., ‘baz’, ‘tingz’).

No reported attempt has been made to test the efficacy of Simplified Wade. However, there are aspects of the proposed system that could potentially work against the advantages of using letters to represent tones. In particular, the choice of obstruent consonants means that a syllable that comprises a consonant followed by a vowel (CV) becomes a differently structured syllable, namely, CVC, and the fact that the added letter provides additional phonemic information might obscure the underlying syllable (e.g., when extracting ‘ba’ from ‘bax’). Moreover, when the underlying syllable ends in ‘ng’ (as many Mandarin words do), the added consonant cannot be integrated with the syllable to form a unified whole. For example, ‘tingv’ is most readily parsed as ‘ting + v’, and ‘tingx’ as ‘ting + x’. As such, it might still be hard to remember which added consonant combines with the syllable to form a particular word since the letter representing the tone is stored separately in memory from the rest of the syllable.

To avoid such problems, the proposal presented here is for a tonal spelling system in which the additional letter representing the tone can be readily integrated with the syllable. The way to achieve this is to place the additional letter immediately after the vowel and to use letters that modify the vowel instead of contributing an additional phoneme. To this end, the letters ‘h’, ‘w’, and ‘y’ are used. Like the other tonal spelling systems, Tone 1 is taken as the basic tone and the other three tones are generated from this. For Tone 2, an ‘h’ is the consonant added after the vowel (e.g., ‘bah’ from ‘ba’, ‘tihng’ from ‘ting’). For Tone 3, it is ‘w’ that is added (e.g., ‘baw’, ‘tiwng’) and for Tone 4, it is ‘y’ (e.g., ‘bay’, ‘tiyng’).

Not only does the use of a letter to represent tone potentially allow for greater integration of the tone with the segmental information, but the assignment of the particular extra letter to its corresponding tone is done in such a way that it adds two additional cues for remembering the relationship between the letter and the tone. First, ‘h’, ‘w’, and ‘y’ are alphabetically ordered when assigned to Tones 2, 3 and 4 respectively and, second, the shape of the letter provides some indication of the change in pitch. That is, ‘h’ has an ascender stroke which is consistent with a rising tone, the strokes of ‘w’ go up and down which is consistent with a dipping tone, and ‘y’ has a descender stroke which is consistent with a falling tone. So, the shape of the letter has the potential to function in the same way as the diacritic used in standard pinyin, but in a way that is integrated with the rest of the syllable. This proposed tonal spelling system will be referred to as “Tones as Alphabetically Formed Tokens” or “TAFT”.

Note that a TAFT syllable might have a strange orthography and be difficult to pronounce (e.g., ‘tiwng’), but it should nevertheless be treated as a single syllable. The strange orthography might even make it more memorable, as has been shown to be the case when children learn words with silent letters in English (e.g., Ehri & Wilce, 1980). For some syllables, the letter-string has a potentially misleading pronunciation if it were spoken as it is written (e.g., ‘baw’) or it might create a real word in the learner’s native language (e.g., ‘bay’ in English). This might not be a problem, however, given that many English speakers seem to have little difficulty learning to pronounce pinyin ‘shi’ with a vowel that is neither /i:/ nor /ai/ (as in the English words ‘ski’ and ‘hi’ respectively), or to not pronounce pinyin ‘you’ as its homographic English word. Again, it is even possible that the orthographic relationship of the pinyin word to the native language might actually make it easier to remember by allowing it to be linked with already stored information.

The present study

The study reported here is a preliminary attempt to establish whether the TAFT system is worth pursuing as a way of enhancing memory for what the appropriate tone should be for Mandarin words that are to be produced. The TAFT system is directly compared to standard pinyin where the tones are represented by diacritics. These two methods of romanized Mandarin were taught to two different groups of native English speakers who knew nothing about the Chinese language prior to the study. Within a single session, they were firstly told about lexical tone and taught the specific tones used in Mandarin, and were then introduced to the romanized transcriptions, with half of the participants being taught TAFT and the other half being taught standard pinyin (the “Pinyin” group). In a paired-learning phase that then followed, each group learnt the Mandarin transcription of each of a set of English words (e.g., ‘LEATHER–wéi’ for the Pinyin group; ‘LEATHER–weih’ for the TAFT group). Finally, the ability to remember the Mandarin words was tested by presenting each of the English words one by one to the participants and asking them to orally produce the Mandarin translation. These responses were recorded and assessed for accuracy by a native Mandarin speaker. In order to measure retention over a longer period of time, the participants returned a week later to perform the same cued-retrieval test.

The reason for adopting the cued-retrieval task was to allow a test of whether learners were able to produce the appropriate tone for a learnt word, even if their pronunciation of that tone was not very native-like. The question being examined was whether those taught using TAFT romanization were able to recall the pronunciation of the Mandarin words better than those taught using standard pinyin.

Method

Participants

The participants were 30 undergraduates from the University of New South Wales in Australia who were given course credit for their participation. All were native English speakers and had no knowledge of Chinese or any other tonal language. They were randomly assigned to either the Pinyin group or the TAFT group, 15 in each with a similar distribution of males and females.

Materials and Procedure

Each session lasted approximately one hour and was conducted by a native Mandarin speaker¹. [Table 1](#) summarizes the procedure, with the Pinyin and TAFT groups differing only after the tones had been taught.

Table 1. *Outline of the procedure (P = “participant”; E = “experimenter”)*

Introduction	P’s were told they would be taught 24 single-syllable Chinese words along with their English meaning and later have to orally produce these Chinese words when given their meaning.
Tone Exposure	P’s were told about tones in Chinese and how they change the meaning of the word. Then, as each Mandarin tone was described with the visually presented labels “Level”, “Rising”, “Dipping”, or “Falling”, E produced the syllable ‘ma’ spoken in the relevant tone and P repeated it.
Tone Learning	P’s copied E using the syllable /o/ spoken with each tone orally labeled as “level”, “rising”, “dipping”, or “falling”. They were then asked to produce the correct tone in response to its spoken label using the syllable /o/, and corrected if wrong. The tone labels were presented in a pseudo-random order across 12 trials with no tone repeated on successive trials. After this, E uttered a syllable in a neutral tone followed by a spoken tone label and the participant had to apply the appropriate tone to that syllable. There were 24 such trials made up of 7 different syllables each with a different vowel, and these were combined in random ways with the different tones (e.g., “wei – falling”; “ming – level”; “qian – falling”; “wei – dipping”). Feedback was given, and any incorrect trials were repeated.
Introduction to Writing Chinese	P’s were told they would be taught a writing system that indicated how to pronounce the syllables that they would be asked to remember: Not Chinese characters, but an alphabetic writing system where tones would be represented.

¹ The invaluable contribution of Lihong Kong to this study is gratefully acknowledged.

	PINYIN GROUP	TAFT GROUP
<p>Introduction to Romanised Spelling</p>	<p>P's were presented with Table 2a and told that the direction of the tone was indicated by a line drawn above the vowel. While being shown the table, P's were told that a level tone was indicated by a straight line, a rising tone by a rising line, a dipping tone by a V-shaped line, and a falling tone by a falling line. E pronounced each example when working through the table.</p> <p>P's were then given 12 new syllables written in pinyin and, for each of these syllables, were asked to say what type of tone it had and then pronounce the syllable with that tone. Errors were corrected.</p> <p>This task was then repeated using the same 12 items, without feedback, and pronunciation accuracy was 93.89%, as recorded by E.</p>	<p>P's were presented with Table 2b and told that no written indication of tone would be given when the tone was level, that an 'h' would be added after the vowel for a rising tone, that a 'w' would be added after the vowel for a dipping tone, and that a 'y' would be added after the vowel for a falling tone. It was pointed out that the added letters were alphabetically organized in the same order that the tones had been taught, and that the shape of the letter gave a clue to the nature of the tone, with 'h' having a line that rises upwards suggesting a rising tone, 'w' having an up-and-down structure suggesting a dipping tone, and 'y' having a line that falls below the line suggesting a falling tone. P's were also told that the only consonants that ever occur at the end of a Chinese syllable are 'n' and 'ng' and, therefore, a final 'w' or 'y' never changes the pronunciation of the vowel as it does in English, being only used as a tone indicator. E pronounced each example when working through the table.</p> <p>P's were then given 12 new syllables written in TAFT and, for each of these syllables, were asked to say what type of tone it had and then pronounce the syllable with that tone. Errors were corrected.</p> <p>This task was then repeated using the same 12 items, without feedback, and pronunciation accuracy was 93.33%, as recorded by E.</p>
<p>Paired Learning</p>	<p>P's were visually presented with 24 syllables paired with their meaning and asked to memorize each pairing so that they could later say aloud the Chinese word. The 24 different meaning-syllable pairings comprised 10 different syllables presented twice or three times with a different tone (e.g., 'LEATHER-wéi', 'DANGER-wēi', 'VISIT-pin', 'ROTTEN-fū'). The English meanings were always single words. P's read the written Chinese syllable out aloud and,</p>	<p>P's were visually presented with 24 syllables paired with their meaning and asked to memorize each pairing so that they could later say aloud the Chinese word. The 24 different meaning-syllable pairings comprised 10 different syllables presented twice or three times with a different tone (e.g., 'LEATHER-weih', 'DANGER-wei', 'VISIT-piyn', 'ROTTEN-fuw'). The English meanings were always single words. P's read the written Chinese</p>

	if the pronunciation was incorrect, E corrected it, and P said it again. Each pairing was presented separately for 5 seconds in a random order. After going through the whole list, it was presented for a second time in a different random order and then for a third time.	syllable out aloud and, if the pronunciation was incorrect, E corrected it, and P said it again. Each pairing was presented separately for 5 seconds in a random order. After going through the whole list, it was presented for a second time in a different random order and then for a third time.
Immediate Cued-Retrieval Test	After a short break, P's were given each meaning alone in a random order and asked to clearly say what they remembered the equivalent Chinese word to be. The spoken responses were recorded and the accuracy subsequently assessed by the native Chinese-speaking E who was blind to the group to which the P belonged.	
Delayed Cued-Retrieval Test	P's returned a week later and were given the same cued-retrieval task, but with the English words presented in a different random order. Performance was assessed in the same way as in the immediate retrieval test.	

Table 2a. Examples given to the Pinyin group.

level	rising	dipping	falling
mā	má	mǎ	mà
mīng	míng	mǐng	mìng
wēi	wéi	wěi	wèi
tān	tán	tǎn	tàn
bō	bó	bǒ	bò
lēng	léng	lěng	lèng

Table 2b. Examples given to the TAFT group.

level	rising	dipping	falling
ma	mah	maw	may
ming	mihng	miwng	miyng
wei	weih	weiw	weiy
tan	tahn	tawn	tayn
bo	boh	bow	boy
leng	lehng	lewng	leyng

Results

Each response was scored as either correct or incorrect separately for the syllable (i.e., the phonemic segments) and the tone. The accuracy scores were then submitted to an ANOVA with the between-group factor being Group (Pinyin vs TAFT), and the within-group factors being Time (immediate vs delayed), and Type of Component (syllable vs tone). Mean scores are found in Table 3.

Table 3. Mean number correct for syllable and tone production comparing the two types of romanization (Max = 24). Standard deviations are in parentheses. Percent correct in italics.

	Immediate		Delayed	
	Syllable	Tone	Syllable	Tone
Pinyin	10.40	11.80	4.80	8.07
	(5.45)	(5.99)	(4.95)	(4.94)
	<i>43.33%</i>	<i>49.17%</i>	<i>20.00%</i>	<i>33.61%</i>
Taft	13.73	15.13	7.93	10.47
	(4.53)	(4.24)	(3.73)	(3.94)
	<i>57.22%</i>	<i>63.06%</i>	<i>33.06%</i>	<i>43.61%</i>
difference	3.33	3.33	3.13	2.40
	<i>13.93%</i>	<i>13.89%</i>	<i>13.06%</i>	<i>10.00%</i>

The Group contrast revealed a significant advantage of TAFT over Pinyin, $F(1, 28) = 4.39$, $p < .05$, $\eta^2 = 0.14$, 95% CI [0.07, 6.03]. For the Time contrast, immediate retrieval was significantly more accurate than delayed retrieval, $F(1, 28) = 45.21$, $p < .001$, $\eta^2 = 0.62$, 95% CI [3.44, 6.46] and, for the Type of Component contrast, scores were higher for tones than for syllables, $F(1, 28) = 14.98$, $p < .001$, $\eta^2 = 0.35$, 95% CI [1.01, 3.29]. There were no interactions with the Group factor (all F 's < 1), but the impact of delay was greater for syllables than for tones, $F(1, 28) = 7.40$, $p < .02$, $\eta^2 = 0.21$, 95% CI [0.19, 1.32].

Discussion

The present study did not address the question of how tonal pronunciation might be improved (e.g., Trísková, 2017) nor how tones might be better identified in speech (e.g., Pelzl, 2019). Instead, the focus was on the question of how Mandarin learners might better remember which tone is the correct one to use for any particular word that is being learnt. It was found that the use of tonal spelling, as exemplified by the TAFT method, allowed learners to recall the Mandarin translation of English words more effectively than the use of pinyin did, both when tested immediately after learning the translations and a week later.

The fact that pinyin represents the tone of a syllable in a different format to the phonemes of that syllable seems to work against the novice learner being able to remember which phoneme-tone combination corresponds to a particular meaning. TAFT overcomes this by integrating the tonal information into the letter-string, hence allowing a single representation to mediate the pronunciation of the word and reduce the cognitive load (see e.g., Sweller, et al., 1990).

In relation to the other main effects, it is no surprise that accuracy was much reduced after a week's delay, while the lower accuracy on the syllables than the tones is simply a reflection of the fact that there are more elements to get wrong in the production of a phoneme-string (i.e., the multiple phonemes) than a single tone. The interaction between Delay and Type of Component follows from this because of the multiple elements that can be lost in the syllable during the delay period as opposed to the single element of the tone.

Theoretical interpretation

It is interesting that the advantage of TAFT over standard pinyin was seen not only in the retention of the tonal information, but also in the retention of the phonemic information. This suggests that the advantage does not arise merely from the greater familiarity with Roman letters than tonal diacritics (e.g., Sweller, 2022), because the letters representing the phonemes are the same in TAFT as in standard pinyin. Instead, it seems likely that TAFT only leads to one unit of information being held in memory (i.e., a single syllable) while pinyin has two (the syllable and the diacritic). As such, TAFT may require fewer resources for the retention of information about the complete pronunciation that includes both the tone and phonemic segments.

One way to consider the difference in lexical representation between TAFT and standard pinyin is that the latter holds the representation of phonemic information separate from the tonal information, hence having the same representation for each tonal version of the syllable (e.g., 'fu' for 'fū', 'fú', 'fǔ', and 'fù'). In contrast, TAFT has four different representations in lexical memory (e.g., 'fu', 'fuh', 'fuw', and 'fuy'), making them less confusable with each other.

These representations might be orthographic, but it is also possible that they take the form of abstract phonological information that is influenced by orthography. Such an idea was proposed by Taft (2006) from a study that asked native speakers of a non-rhotic dialect of English (Australian) to decide whether pseudowords like CAWN were homophonic with a real word. In a non-rhotic dialect, CAWN is homophonic with the word 'corn' (i.e., /kɔ:n/), but the participants had considerable difficulty realizing this when the task was performed silently. Only when the participants read the pseudowords out loud before responding did they become aware of the homophony. Taft (2006) argued that the homophony of CAWN with 'corn' was not recognized under silent conditions because the underlying phonological representation of 'corn' includes the 'r', as manifested in the orthography, and that the pronunciation /kɔ:n/ only emerges when the surface form is overtly generated from the underlying representation.

In the same way, it is possible that the phonological representations set up in Mandarin are influenced by the spelling of the romanized script through which it is learnt. If so, the underlying phonological representation of a syllable is different for the four tones when learnt via the TAFT method, but the same when learnt via standard pinyin. Moreover, the appropriate surface pronunciation appears to be readily generated from the underlying TAFT representation given that participants in that group were able to avoid pronouncing the letter-string as though it was English. For example, no-one thought that 'ba' with a dipping tone (i.e., 'baw') was pronounced in the same way as it would be in English (i.e., /bɑ:/).

Pedagogical implications

It is important to note that, because of its small scale, the present study should only be seen as a preliminary attempt to demonstrate the advantage of the TAFT method over standard pinyin. If the findings are upheld in future research, however, the implication would be that the use of romanization as a mediating tool for teaching Mandarin as a foreign language might be expedited if tonal information were represented by letters rather than tonal diacritics, at least for learners whose L1 does not use diacritics. The two most commonly used indicators of tone are the diacritic lines tested here or an added digit to indicate the tone number such as ‘fu(1)’, ‘fu(2)’, etc. While the latter has not yet been tested against tonal spelling, there is no reason to believe that it would be any better than the use of diacritics since it too does not allow the tone and phonemic segments to form an integrated unit.

The TAFT method provides a systematic and logical approach to tonal spelling and the results of the present research are encouraging with regard to its adoption in courses that teach Mandarin as a foreign language, especially when the students are familiar with Roman script. Whether or not it is more effective than pinyin for those whose native language uses diacritics or does not use the Roman alphabet is something that remains to be seen.

Multisyllabic words

If TAFT were to be adopted as a teaching tool, consideration needs to be given to the method by which words with more than one syllable are handled, a situation that is very frequent in Chinese. In standard pinyin, multisyllabic words are typically written without a space between syllables, but TAFT faces a potential problem with ambiguity when there is a vowel at both the end of the first syllable and the beginning of the second. For example, the TAFT version of the two-syllable word ‘tiàn’ (meaning “a proposal”) is ‘tihayn’, but this could also correspond to ‘tīhàn’. For this reason, a TAFT transcription might benefit from having the individual syllables of a multisyllabic word separated from each other with a space. A hyphen might even be preferable since it would keep the syllables of the word together.

Neutral tones

It should also be noted that, very occasionally, word-final Mandarin syllables have a neutral tone, which is short and lacking dynamics. This neutral tone is represented in standard pinyin by simply not including a tone marker (e.g., the second syllable of ‘shénme’ meaning “what?”). If TAFT were to also represent the neutral tone without a tone marker, there would be no way to differentiate it from Tone 1. This was also a problem facing both the GR and Simplified Wade tonal spellings because they also had no marker for Tone 1. So, the solution they adopted can also be incorporated into TAFT, namely, to place a dot prior to the neutral syllable (e.g., ‘shehn .me’). Another possibility is to capitalize the vowel (e.g., ‘shehn mE’). It should be said, however, that there are only a few cases of neutral tones in Mandarin and the words in which they occur tend to be very common. So, it is also possible to make no change to the syllable and simply point out to the learner that these particular cases take a neutral tone rather than Tone 1.

Conclusions and limitations

The present research provides an indication that a systematic tonal spelling system, as exemplified by the TAFT method, might be a better way to mediate the learning of the appropriate pronunciation of Mandarin words. The results show better retention not only of the correct tone of newly learnt words, but of the correct segmental information as well. The implication is that the acquisition of Mandarin vocabulary might be facilitated through the use of a system like TAFT where pronunciations are represented alphabetically.

Having said this, however, further research will be necessary to confirm the utility of the TAFT method given the small number of participants per group in the present study and the use of a laboratory context. It might be informative to follow something like the procedure of McGinnis (1997) by comparing students on their knowledge of the appropriate pronunciation of words after having been taught either via pinyin or via TAFT during a year-long introductory Mandarin program. In addition, the performance of such students should be evaluated by more than one native Mandarin speaker to ensure greater reliability, unlike the present study.

Finally, it should be noted that the widespread use of standard pinyin in such things as learning resources and digital communication means that, even if TAFT were adopted by Chinese learners, the way it corresponds to the diacritic representation of tone would also need to be learnt in order to read and generate pinyin. However, given the one-to-one mapping between the additional letters used in TAFT and the pinyin diacritics, there should be little difficulty in learning the system used for reading and typing pinyin in association with the TAFT system that was learnt for the purposes of remembering the appropriate tone of any particular word.

References

- O. B. Anderson. (1967). A simple and consistent mode of transcription for standard Chinese, Cantonese, and Siamese. *Libri*, 17, 86-94. <https://doi.org/10.1515/libr.1967.17.1-4.86>
- Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, 15, 254-262. <https://doi.org/10.1016/j.tics.2011.04.003>
- Ehri, L. C., & Wilce, L. S. (1980). The influence of orthography on readers' conceptualization of the phonemic structure of words. *Applied Psycholinguistics*, 1, 371-385. <https://doi.org/10.1017/S0142716400009802>
- Godfroid, A., Lin, C. H., & Ryu, C. (2017). Hearing and seeing tone through color: An efficacy study of web-based, multimodal Chinese tone perception training. *Language Learning*, 67, 819–857. <https://doi.org/10.1111/lang.12246>
- Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience*, 27, 649-677. <https://doi.org/10.1146/annurev.neuro.27.070203.144220>
- Liu, Y., Wang, M., Perfetti, C. A., Brubaker, B., Wu, S., & MacWhinney, B. (2011). Learning a tonal language by attending to the tone: An in Vivo experiment. *Language Learning*, 61, 1119–1141. <https://doi.org/10.1111/j.1467-9922.2011.00673.x>
- Kratochvíl, P. (1968). *The Chinese language today*. Hutchinson. ISBN 0-09-084651-6
- McGinnis, S. (1997). Tonal spelling versus diacritics for teaching pronunciation of Mandarin Chinese. *The Modern Language Journal*, 81, 228-236. <https://doi.org/10.2307/328789>
- Pelzl, E. (2019). What makes second language perception of Mandarin tones hard? A non-technical review of evidence from psycholinguistic research. *Chinese as a Second Language*, 54, 51–78. <https://doi.org/10.1075/csl.18009.pel>
- Showalter, C. E., & Hayes-Harb, R. (2013). Unfamiliar orthographic information and second language word learning: A novel lexicon study. *Second Language Research*, 29, 185-200. <https://doi.org/10.1177/0267658313480154>
- Sweller, J. (2022). The role of evolutionary psychology in our understanding of human cognition: Consequences for cognitive load theory and instructional procedures. *Educational Psychology Review*, 34, 2229–2241. <https://doi.org/10.1007/s10648-021-09647-0>

- Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology: General*, 119, 176–192. <https://doi.org/10.1037/0096-3445.119.2.176>
- Taft, M. (2006). Orthographically influenced abstract phonological representation: Evidence from non-rhotic speakers. *Journal of Psycholinguistic Research*, 35, 67-78. <https://doi.org/10.1007/s10936-005-9004-5>
- Třísková, H. (2017). Acquiring and teaching Chinese pronunciation. In *Explorations into Chinese as a second language* (pp. 3–30). Springer

About the Author

Marcus Taft is both a Distinguished Professor at the Centre for the Cognitive Science of Language, Beijing Language and Culture University, PR China, and an Emeritus Professor at the School of Psychology, UNSW Sydney, Australia. He has published numerous papers on the topic of lexical processes in reading across different languages. The idea for the present study arose from his own quite unsuccessful attempts to learn Mandarin.

E-mail: m.taft@unsw.edu.au