中央研究院歷史語言研究所集刊
第六十八本，第四分
出版日期：民國八十六年十二月

A Phonetic Study of Tsou

Richard Wright* and Peter Ladefoged**

This paper examines the phonetic characteristics of the Tfuea dialect of Tsou, an Austronesian language spoken in Southern Taiwan. The authors employ both acoustic and auditory analyses, as such it represents the first instrumental study of Tsou. As the consonant and vowel inventories of the language are examined in phonetic detail, several points of disagreement in previous descriptions of the language are cleared up. The analyses include vowel formant measures, consonant voicing and VOT by place of articulation, and intrinsic pitch of vowels. In addition to the segmental description, there is a preliminary investigation of the consonant clusters many of which are only rarely attested in the world's languages.[1]

Keywords: Tsou, Acoustic Phonetics, Consonant inventory, Vowel inventory, Consonant clusters

---

* Indiana University

** University of California, Los Angeles

[1] This research was supported by the National Science Council of Taiwan and the Academia Sinica, Taiwan, and by NSF. We are grateful to Elizabeth Zeitoun for fieldwork assistance, Paul Li for his encouragement and valuable comments, and to Chiu-yu Tseng for her efficient arrangements. Thanks also go to two an anonymous reviewers. Special thanks are due to Mo'o Muknana who hosted our visit in Pnguu and to all the villagers who cooperated so enthusiastically.

Richard Wright & Peter Ladefoged

Introduction

Tsou is an Austronesian language spoken by approximately 3000 people on Mount Ali in Central Taiwan. Historically it was made up of four dialects, Tapangu, Tfuea, Luhtu (Mamahavana) and Iimucu, the last of which is extinct (Tung, 1964). In the three remaining dialects there is no significant grammatical and only marginal phonological variation. Aspects of Tsou have been described in various forms in several previous studies: Nevskij (1935), Ogawa & Asai (1935), Wei, Yu & Lin (1952), Lin (1955), Tung (1964), Starosta (1974), Tsuchida (1972, 1976), Ho (1976), Li (1979) and Szakos (1994). Of these Tung's descriptive study is the most comprehensive, but those by Li (1979), Tsuchida (1976) and Szakos (1994) provide important additions and corrections to Tung's description of the language. In a recent dissertation, Wright (1996) investigates Tfuea consonant clusters in phonetic detail.

The present study is based on recordings made in 1993 of 13 speakers, 8 males and 5 females, of the Tfuea, **'tfuja**, dialect as spoken in Pnguu village. It reports on the major phonetic characteristics of the Tsou language as determined by both acoustic and auditory analysis, and as such represents the first instrumental study of the phonetic structures of Tsou.

Phonological overview

Tsou has six vowels: i, e, ɨ, u, o, a. According to Tung (1964), there is some dialectal variation associated with i, in that Tapangu i corresponds to z or iz in Tfuea and Luhtu in certain environments. Another variation, Luhtu r corresponding to Tapangu and Tfuea j may no longer be viable, since Li (1979) could find only one Luhtu speaker, over 70, who still preserved it.

Tung describes long vowels as sequences of like vowels as most of the apparent long vowels occur as a result of affixation, reduplication or vowel lengthening. He

also describes long vowels as heterosyllabic and cites evidence from stress assignment and reduplication. Szakos presents a list of minimal pairs illustrating length contrast for the vowels i, u, a. However, his examples contain long vowels that are the result of morphological processes. In light of modern theories of metrical stress (eg Hayes, 1995) and morphophonology (McCarthy and Prince, ms), the reduplication and stress assignment in Tsou can be seen as weight dependent processes that count morae rather than syllables. One reason for viewing many of the 'disyllabic' vowels as long is that there is no hiatus between adjacent homorganic vowels. The very processes that Tung cites as evidence for treating long vowels as heterosyllabic can also be seen as evidence in favor of long vowels being treated as tautosyllabic. In the end, Tung's observation that there is no phonemic length contrast appears accurate.

Tsou was described by Tung (1964) as having two 'non-syllabic' vowels that are the counterparts of e and o. However, Ho (1976) and Li (1979) describe the 'non-syllabic' vowels as the glides j and w, which seems to us a sensible proposal from a phonological point of view, especially given the stress pattern and the process of reduplication seen in the language. It should be noted, however, that w has a limited distribution and that in the present study, there is no systematic data available for its analysis. Tung's analysis is reflected in the orthography which, along with a phonetic transcription, will be used in this paper. Thus the orthographic version of the name of the dialect is 'Tfuea' which in our transcription is **tfuja**.

In addition to the two glides, the Tfuea dialect has the following 15 consonant sounds: ɓ, ɗ, p, t, k, ʔ, ts, f, s, h, v, z, n, m, ŋ. Tung (1964) notes that Luhtu has r where Tfuea has j; however even at the time of his study, Luhtu speakers freely substituted j for r. Tsuchida, and later Li, noted that at that time only the very elderly speakers maintained the r. An additional consonant series, pulmonic ingressive fricatives, has been said to occur in the Tfuea dialect spoken in Pnguu (Fuller 1990), but Ladefoged and Zeitoun (1993) note that all of the 14 Pnguu speakers they investigated used an egressive airstream to produce the fricatives described by Fuller.

One of the most striking features of Tsou Phonology is the wide range of combinatorial possibilities of the 15 consonant sounds. Sequences of more than two consonants are forbidden, but a large number of unusual two consonant clusters do occur. These clusters will be discussed below.

Vowels

For five of the six vowels the data consisted of two examples of each vowel in a stressed syllable after a bilabial stop, and two further examples in a stressed syllable after an alveolar stop. This was not possible for the sixth vowel, ɨ, for which only two examples from a stressed syllable following an alveolar stop were available in the data which we had recorded. Examples of the vowels are given in Table 1

VOWEL	ORTHOGRAPHY	IPA	GLOSS
i	*pitu*	pitu	'seven'
e	*pepe*	pepe	'sky, heaven'
a	*pa'ti*	paʔti	'to show'
o	*po'e*	poʔe	'to pray in making wine'
u	*tufku*	tufku	'to wash clothes'
ɨ	*tɨ'sɨ*	tɨʔsɨ	'arrow'

Table 1: Words illustrating the short vowels of Tsou

The word list used in making the recordings is given in the appendix. Each word was produced once by each speaker. While minimal sets were difficult to find, the data were selected to avoid consonant environments that would cause allophonic variation. Formant and fundamental frequency measurements were taken in stressed syllables.

Measurements of the frequencies of the first, second and third formants (F1, F2, and F3 respectively) were taken, using spectrographic displays and superimposed FFT and LPC

spectra on the KAY CSL system. In order to maximize the accuracy of the spectral analysis, speech sampled at 10 kHz was used in the instrumental analysis of vowels (see Ladefoged, 1993). The LPC calculation used a 20 ms frame with pre-emphasis applied to the signal prior to calculation to make it easier to see the higher formants. In general, a 14th order filter was applied in analyzing the male speech and a 12th order filter was applied in analyzing the female speech. When necessary the filter order was adjusted down to accommodate the higher fundamental frequency of the speech in certain utterances for some speakers. To insure the accuracy of the LPC measurements, a simultaneous 512 point FFT calculation was made and the results were overlaid on the LPC results. These measures were also checked against measures taken from a spectrogram and against auditory impressions. The mean formant frequencies are shown in Table 2.

	female speakers			male speakers		
vowel	**F1**	**F2**	**F3**	**F1**	**F2**	**F3**
i	435	2599	3168	369	2157	2864
e	516	2328	2923	499	2032	2597
a	975	1581	2952	805	1385	2629
o	560	914	2963	539	983	2576
u	468	1118	3158	407	917	2464
ɨ	501	1756	3012	429	1441	2454

Table 2: Mean formant frequencies for male and female speakers

Formant frequencies are plotted in Figure 1 (males) and Figure 2 (females). In the figures F1, roughly equivalent to vowel height, is plotted against F2' – F1 which is roughly equivalent to the backness of the vowel. F2' is a weighted average of F1, F2 and F3 calculated with the formula given by Fant (1973: 52): F2' = F2 + (F3 – F2) (F2 – F1) / 2 (F3 – F1). The axes are scaled using the Bark scale and labeled in Hz. The ellipses around the points for each vowel have a radius of two standard deviations from the mean.

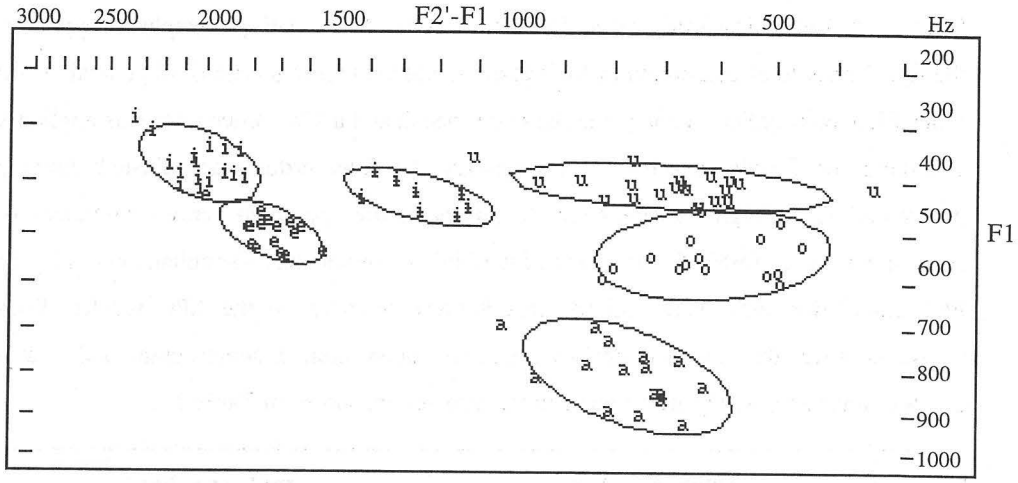


Figure 1: Formant plot for male speakers F2' – F1 by F1

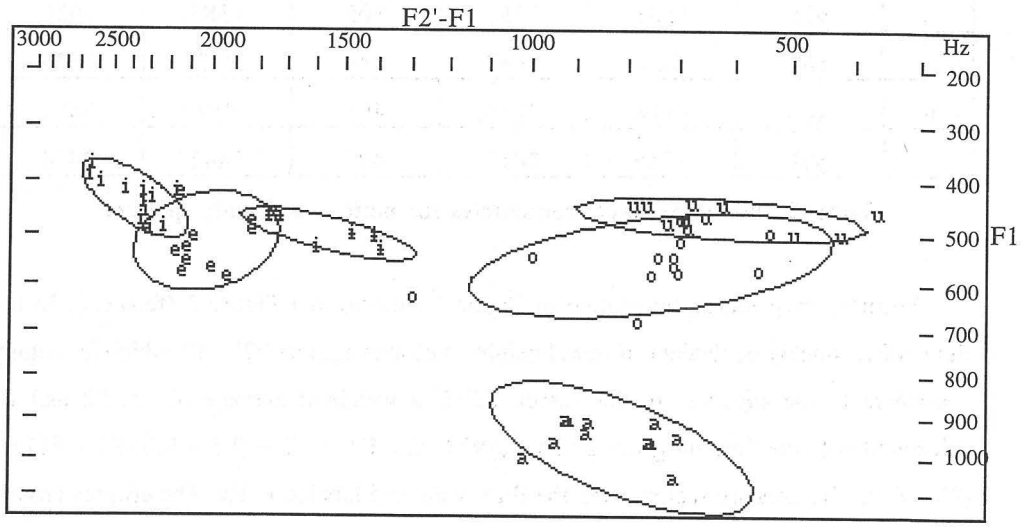


Figure 2: Formant plot for female speakers F2' – F1

Based on acoustic data and auditory impressions the following description of Tsou vowels can be made:

i high front unrounded
e midhigh front unrounded
ɨ high central unrounded
u high back rounded
o midback rounded
a low back unrounded

The vowel e has been described by Tsuchida as 'lower mid-front' corresponding to ə. It is plain from the formant plots that the midfront e in Tsou is higher than what is traditionally transcribed as ə, and is much closer to i than to a. Tung describes ɨ as high central unrounded, but chooses to transcribe it as ㄔ because his impression is that it is closer to u than to i and is slightly rounded. Observations of the speakers' lips during production revealed no evidence of rounding. It seems to us that it is a high central vowel as described by Li and Tsuchida, and is more accurately transcribed as ɨ, in accord with Fuller (1990) and Szakos.

The most notable aspect of the vowel space in Tsou is the centralizing tendency, particularly in the female speakers. The centralization was greatest when the vowel was flanked by voiceless coronals and is most significant in the back vowels. There was a short transition into the vowel and a longer schwa-like transition out of the vowel and into the following consonant. This was particularly pronounced in productions of o by one of the female speakers, as is reflected in the wide scatter in the formant plot.

In addition to formant measures, inherent pitch measures were taken for each vowel. Fundamental frequency (F0) was measured from a narrow band spectrographic display that was linked to simultaneous wide band spectrographic, waveform and amplitude contour displays. F0 measurements were taken in a stressed syllable at the point that corresponded to the F0 peak for that syllable. The F0 means are plotted with standard deviations in Figure 3.

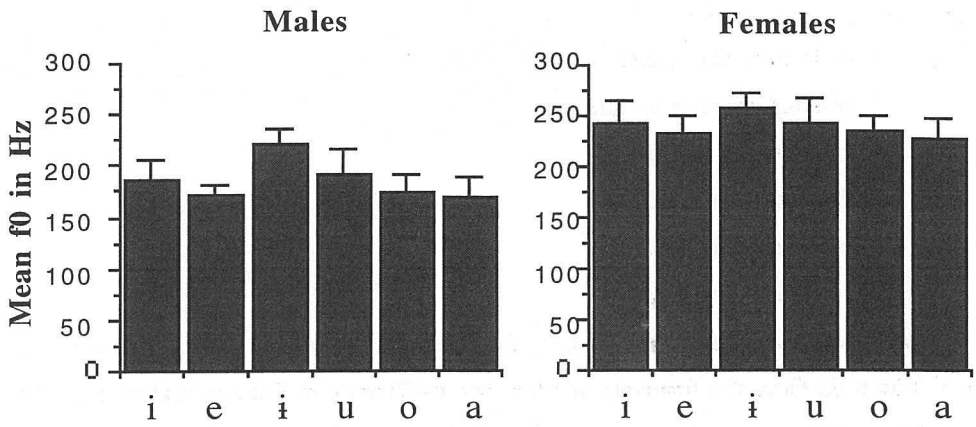


Figure 3: Mean and standard deviations for male and female F0 values for vowels

Surprisingly, markedly higher F0 values were found for the high central vowel than for the other high vowels despite the fact that on the formant plots this vowel is lower than either i or u. The males had larger differences in F0 between vowels than the females. Normalized F0 and F1 were statistically correlated to determine the reliability of the effect of vowel height on intrinsic pitch. The normalized values in the correlation were the individual deviation from the overall mean for F1 and F0. As is shown in Table 3, a highly significant but weak correlation was found between the deviation from mean of F1 and F0. This result is in agreement with the general finding that vowel height is correlated with fundamental frequency such that high vowels have an intrinsic pitch that is higher than low vowels.

Correlation	P-Value	95% Lower	95% Upper
-.333	<.0001	-.465	-.187

Table 3: Correlation of vowel height (F1) and F0

Consonants

As was noted above, the Tfuea dialect of Tsou has the 17 consonants listed in Table 4 below. Both auditory and acoustic analysis were used in arriving at the following description of Tsou consonants. In the recordings there was one repetition of each consonant in onset position before both i and a, except for the glides j, w. Words containing the consonants were digitized at either 10 kHz (sonorants) or 20 kHz (obstruents) and analyzed using simultaneous waveform, amplitude and spectrographic displays on the KAY CSL. The wordlist used to record the consonant tokens is found in the appendix.

	BILABIAL	LABIO-DENTAL	ALVEOLAR	PALATAL	VELAR	GLOTTAL
PLOSIVE	p		t		k	ʔ
IMPLOSIVE	ɓ		ɗ			
AFFRICATE			ts			
FRICATIVE		f v	s z			h
NASAL	m		n		ŋ	
APPROXIMANT	w			j		

Table 4: Chart of Tsou consonants

Plosives:

Aspiration is not contrastive in Tsou. While voiceless stops can appear aspirated, Tung cites both distributional and morphophonemic evidence for treating these surface forms as clusters. The strongest distributional evidence is the fact that h clusters freely as either the first or the second member of a cluster but blocks further clustering. Tsou permits maximal clusters of two consonants; clusters of h and two or

more consonants are not permitted. Consonant clusters are discussed below. In certain inflectional processes, the first member of a consonant-h cluster can change or in other processes be removed leaving behind the h. The following example, the verb 'to trade', is taken from Tung (1964:11):

mhino (actor focus), phini (benefactor focus).

Description of the voiceless plosives is straightforward: p is a voiceless unaspirated bilabial, t is a voiceless unaspirated alveolar, k is a voiceless unaspirated velar. The status of the glottal stop as a consonant is supported by distributional evidence. Like the glottal fricative, h, it is found with the same distribution as the other consonants. Both of these consonants can occur in intervocalic position, in consonant clusters as the first or second member. Like h it blocks further clustering. Moreover, Szakos (p. 22) presents some examples that provide evidence to corroborate Tsuchida's brief note describing a contrast between a word initial glottal stop and a word initial onsetless vowel. Such examples further support the phonemic status of the glottal stop in Tsou. As Szakos's work was published subsequent to the data collection for this study, no instrumental study of the contrast was undertaken.

The voiced bilabial stop has been described as a preglottalized voiced bilabial plosive by Tung, Li, and Tsuchida, as a preglottalized voiced ejective by Szakos, and as a voiced bilabial implosive by Fuller. Spectrographic analysis of ɓ in word onset shows no creakiness or glottalization during voiced closure or in the onset of the vowel. Visual observations of lowering of the speaker's larynx during closure indicates that the voiced bilabial is in fact implosive. Inspection of the waveform and spectrogram reveals that voicing during closure is characteristic of that seen in implosives (Lindau, 1984) in that voicing is strong and there is an amplitude increase in the latter part of the closure. An expanded view of the waveform showing voicing during ɓ closure is shown in Figure 4.

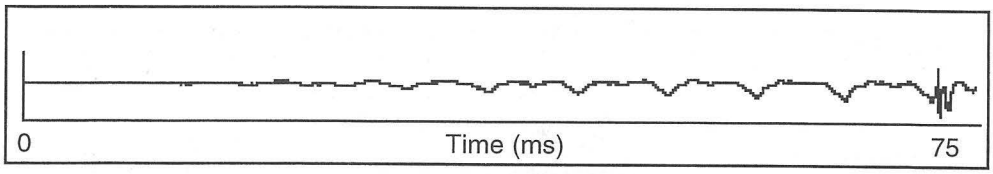


Figure 4: Expanded waveform showing increase in amplitude during ɓ closure voicing characteristic of implosives.

The voiced alveolar stop has been described variously; Tung describes it as a preglottalized alveolar lateral l, Tsuchida describes it as a voiced alveolar stop d and Fuller reports hearing it as an implosive ɗ. Auditory impressions, acoustic measures and observation of the speakers larynx indicate that it is implosive. Like ɓ, it shows an increase in amplitude during the closure voicing, as shown in Figure 5.

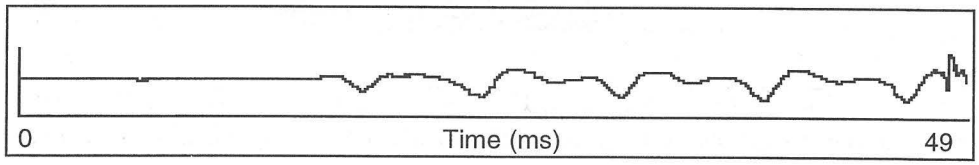


Figure 5: Expanded waveform showing increase in amplitude during ɗ closure voicing characteristic of implosives.

Both ɓ and ɗ may be produced with preglottalization which is not unexpected in implosives due to the nature of their articulation. However, it is worthy of note that the glottal constriction is more marked than is usually seen in examples in the literature and often results in a cessation of voicing prior to the stop's voicing. The preglottalization is not evident in word initial position, but in intervocalic position it is clear as can be seen in Figure 6. What is unusual is that one third of the speakers produce the voiced alveolar implosive with a lateral release and all of the speakers produce a lateral approximant allophone l before the low vowel a. Between vowels the lateral allophone is clearly accompanied by preglottalization. The lateral approximant in word onset is shown in Figure 7 below. The lateral characteristics are probably the source of some of the disagreement seen in past descriptions. Szakos points out that there is a lateral flap before high and mid vowels in a very limited set of lexical items. As all of the words which exhibit the lateral flap before high and mid vowels are fairly recent loans from Japanese, they have been excluded from this study.

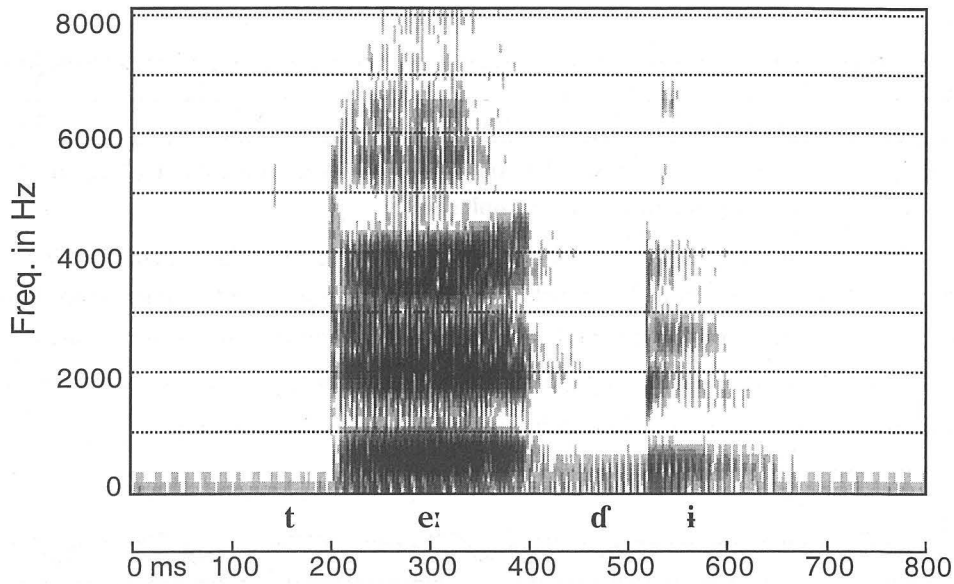


Figure 6: Spectrogram illustrating preglottalization of ɖ in the word teːɖi 'to arrive on time'

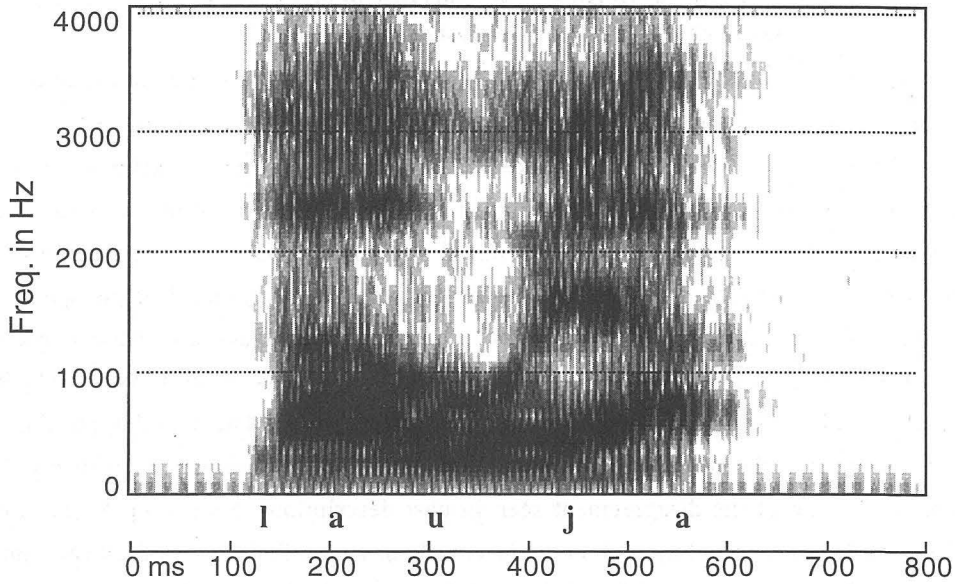


Figure 7: Example spectrogram showing the lateral approximate allophone of ɖ in the word ɖauja [lauja] 'maple'

In addition to the spectrographic analysis, voice onset time (VOT) measures were taken for the voiceless unaspirated plosives. The interval between the release burst and the first glottal pulse was measured on simultaneous waveform and spectrographic displays. The measurements were submitted to an analysis of variance with VOT as the dependent variable and consonant and speaker as independent variables. The results ($F[2,21]=47.15$, $p<.0001$) indicated that there was a significant difference in VOT between the consonants. The mean VOT following the three voiceless plosives is shown in Figure 8 below plotted with error bars. The results agree with the general observation that the further back a consonant closure is made, the longer the duration of the VOT. (Fischer-Jørgensen, 1954).

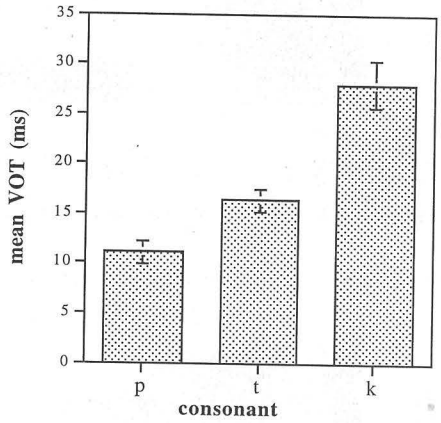


Figure 8: Mean VOT in ms following p, t, k

Affricates:

There is a single affricate in Tsou, the alveolar: ts, illustrated by a spectrogram in Figure 9. Its frication is characterized by a concentration of energy between approximately 5 kHz and 6 kHz It is palatalized before the front vowels i and e, being realized as the palato-alveolar affricate tʃ, with a wider distribution of energy that spreads between approximately 3.5 kHz and 5 kHz, illustrated by a spectrogram in Figure 10. ts is a single phonological unit rather than a cluster of t + s, as is shown by its combinatorial possibilities, to be discussed below.

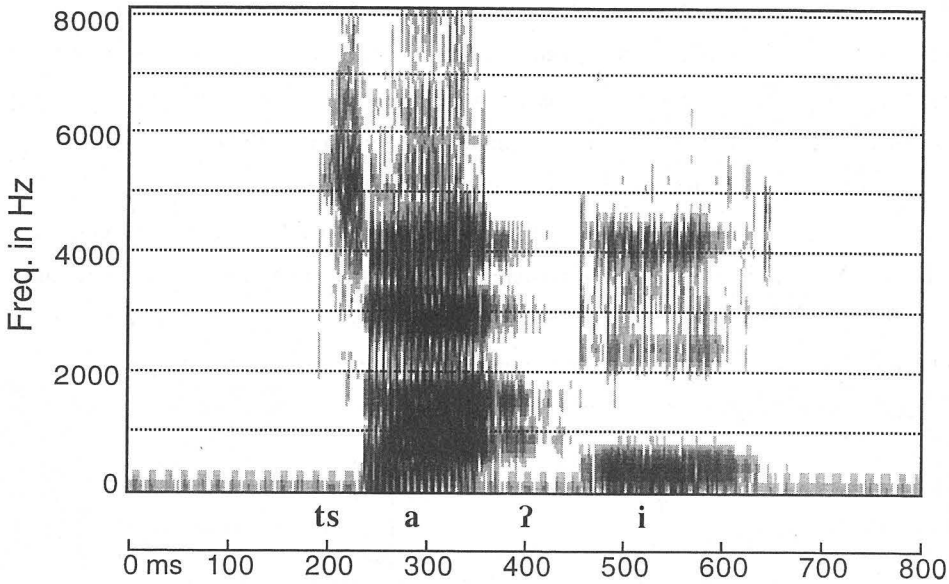


Figure 9: Spectrogram illustrating the alveolar affricate ts **in the word** tsaʔi **'dirty'**

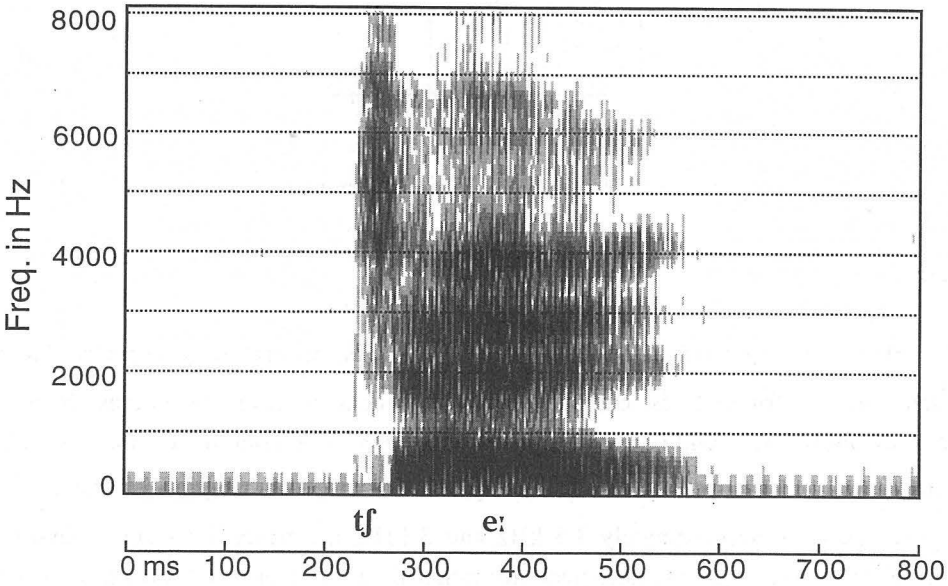


Figure 10: Spectrogram illustrating the palatalized allophone tʃ **in the word** tʃej **'dream'**

Fricatives:

There are five fricatives in the Tfuea dialect of Tsou: voiceless and voiced labiodental f, v; alveolar s, z; and the glottal fricative h. The frication energy of the voiceless labiodental fricative f has low intensity and is widely spread over the spectrum as is typical in labiodentals, illustrated in Figure 11. Some speakers produce the frication with a more concentrated energy particularly in a cluster, such as ft, illustrated in Figure 12.

The voiceless alveolar fricative, s, is illustrated in Figure 13. The voiceless alveolar fricative is characterized by a high frequency concentration of energy. Like the alveolar affricate, the alveolar fricatives are heavily palatalized before front vowels. The palatalized form, ʃ, is illustrated in the spectrogram in Figure 14.

Tsou h has been described by Tsuchida as a velar fricative and by Tung as a glottal fricative which varies in quality depending on the environment. From a phonetic point of view the data in this study, illustrated in Figure 15, support Tung's description. As is common to h, the glottal fricative in Tsou often shows concentration of energy in the region of F2, and sometimes F3, of the following vowel, and takes on the qualities of its vocalic environment. Tsuchida's description may have been influenced by the fact that there is significant coarticulation between h and the high central unrounded vowel, ɨ. In this environment h is often realized with spectral qualities that are characteristic of a velar fricative, an example of which is shown in Figure 16. There may be phonological support for the view that h is a velar fricative, or at least not a glottal one. Tsou has a prohibition against two obstruents at the same place of articulation forming a cluster. But the clusters hʔ and ʔh do occur, as illustrated in Figure 15. However, as many phonologists regard prohibitions against homorganic clusters as applying only to oral articulations, glottal articulations should be exempted from such a prohibition. In addition, Wright (1996) and Szakos (1995) have found hk clusters. It should be noted here that in the Luhtu dialect both Li and Tsuchida analyze this fricative as phonetically a velar x, but phonologically as a glottal fricative. As the acoustic analysis in this study is based exclusively on recordings of the Tfuea dialect, the authors must defer to Li and Tsuchida's analysis of Luhtu.

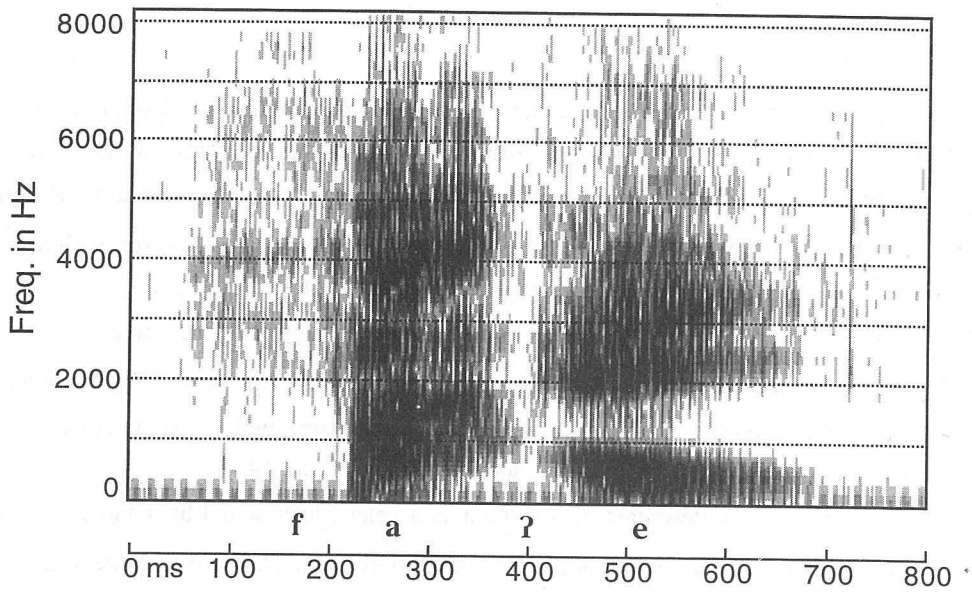


Figure 11: Example spectrogram: f with broad spectrum energy during frication in faʔe **'a man's name'**

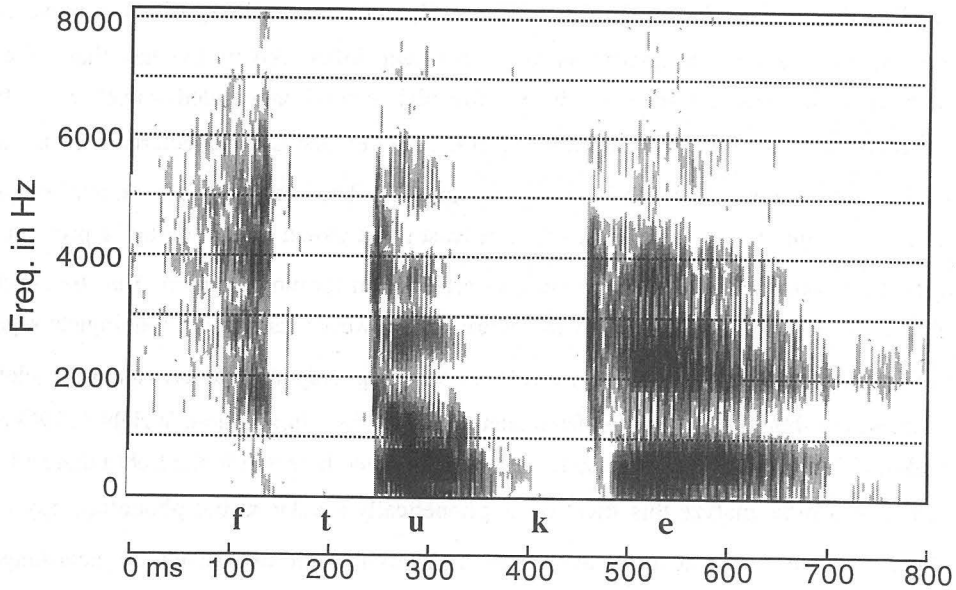


Figure 12: Example spectrogram: f in a cluster ftuke **'to slouch'**

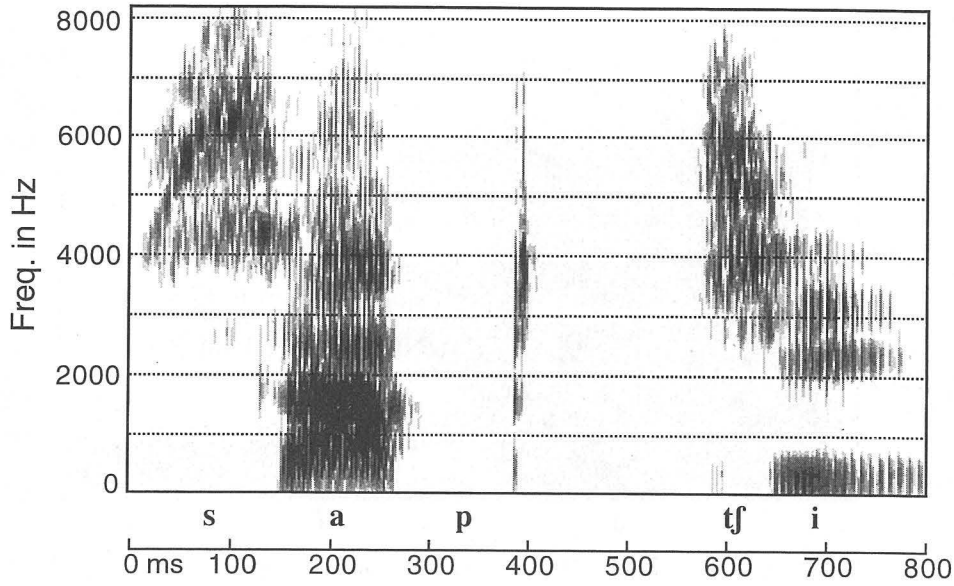


Figure 13: Example spectrogram: s in saptʃi **'face'**

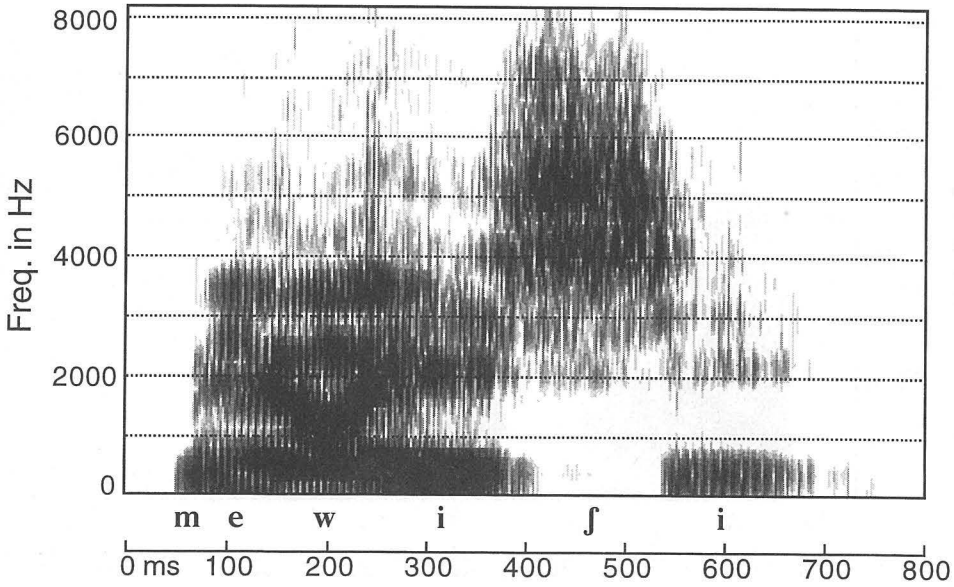


Figure 14: Example spectrogram: ʃ in mewiʃi **'large'**

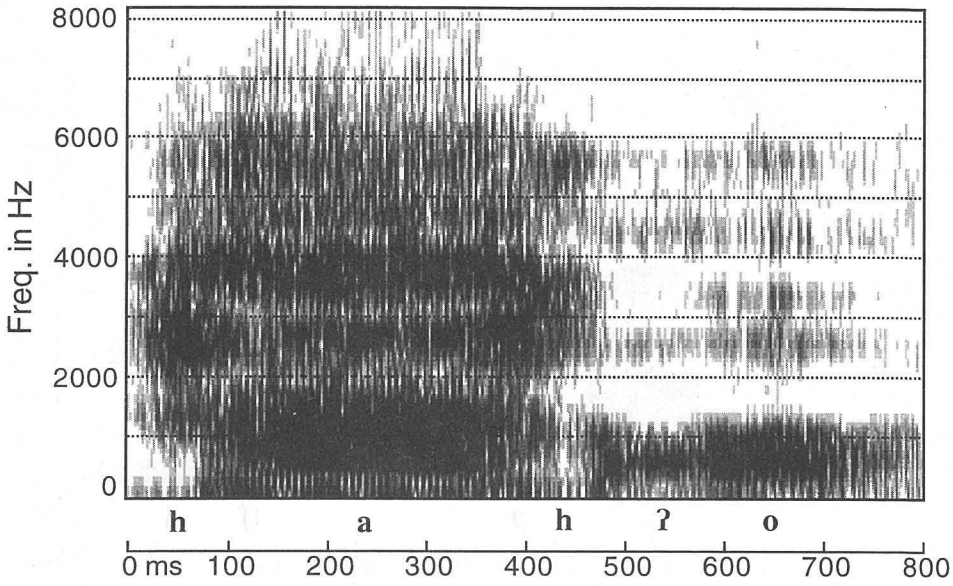


Figure 15: Example spectrogram: h **in** hahʔo **'friends'**

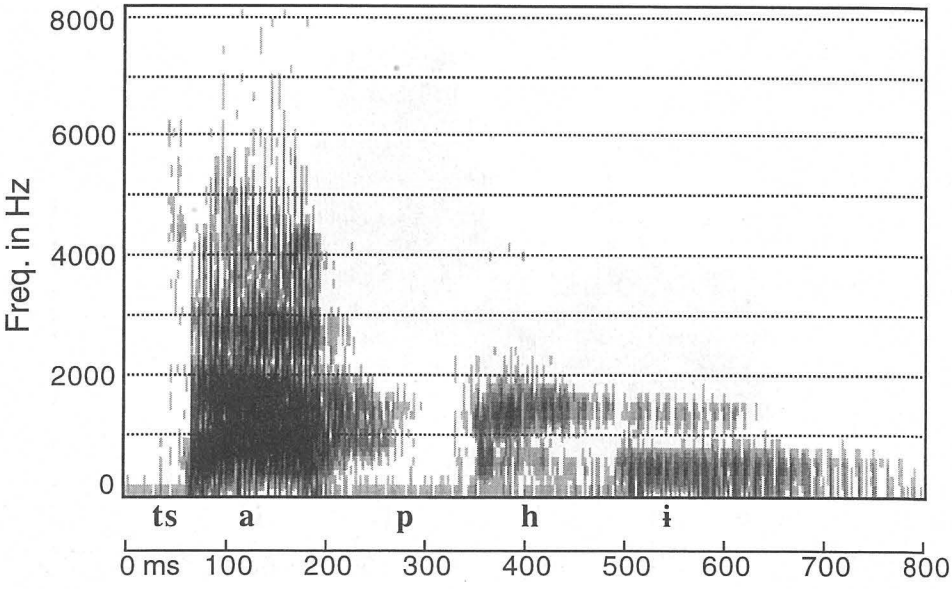


Figure 16: Example spectrogram: h **in** tsaphi̠ **'palm, sole'**

Nasals:

The nasal consonants have three places of articulation paralleling those of the voiceless unaspirated plosives: bilabial m, alveolar n, and velar ŋ. The most remarkable thing about the nasals is that, with the possible exception of h and ʔ, if these are both considered to be glottal consonants, they are the only segments that are allowed to form homorganic clusters with a following obstruent.

Approximants:

Tung describes two 'consonantal vowels' which he transcribes ĕ and ŏ in his discussion of vowels but which he fails to differentiate from e and o elsewhere in his text. Ho and Li have suggested treating the consonantal vowels as j and w. The rationale is that phonological processes such as stress assignment and reduplication are systematic only if the segments in question are consonantal. For example, certain inflectional processes involve reduplication of the onset and first vowel, or first mora, of the base. The following examples of reduplication and stress assignment are taken from Tung, using his transcription as the orthography.

Orthography	IPA	Gloss
ex: *cofkoea*	tsofkoja	'clean'
cocofkoea	tsotsofkoja	'very clean'

When ĕ is found at the onset of a word, it is reduplicated as if it were a consonant.

ex: *eoskɨ*	joskɨ	'fish'
eoeoskɨ	jojoskɨ	'fish' pl.

When ĕ is found as part of a syllable preceded by a consonant and followed by a vowel, it is treated as part of a consonant onset cluster and it is reduplicated together with the preceding consonant and first vowel of the base.

ex: *beahci*	bjahtʃi	'fruit'
beabeahci	bjabjahtʃi	'fruits'

If ĕ is treated as a syllabic vowel, the rules for reduplication become very complicated indeed. If, on the other hand it is treated as a glide, rules for reduplication are straightforward: reduplicate the onset and first vowel of the base. Stated in terms of prosodic morphology, this is a process of light syllable reduplication. The syllabic front midvowel e may itself form a syllable nucleus as is evident in forms such as pepe 'sky'. Furthermore, e is reduplicated in words in which it makes up the nucleus of the first syllable:

ex: *evi*	evi	'tree'
eevi	eevi	'trees'

This type of evidence from reduplication is not available for ŏ, but evidence from stress assignment does point to treating both ŏ and ĕ as glides, distinct from the front and back midvowels. In Tsou, stress generally falls on the penultimate syllable.

ex:

óko	ˈoko	'child'
okósi	okˈosi	'his child'
mʔúmʔu	ˈmʔumʔu	'hair'
eamʔumʔúa	eamʔuˈmʔua	'hair'
úa	ˈua	'deer'
iunzóu	iunzˈou	'to burn'
ciuéi	tʃiuˈei	'to pierce the ears'

The first two pairs of examples show that stress is mobile and occurs on the penultimate syllable; when a suffix is added to the word, the stress moves to the syllable that was final in the non-suffixed form. The fact that stress is mobile in Tsou is evidence in favor of an analysis, such as that by Tung, Li and Ho, in which Tsou stress is quantity sensitive and not assigned underlyingly. The last four examples above show that in a series of vowels, the stress falls on the second to last vowel. As can be seen both o and e are treated as syllabic as far as these forms are concerned since both receive stress. In other forms, however, what have been transcribed by Tung as o and e are treated as non-syllabic.

ex:

kóeu	ˈkoju	'ear'
koeúsi	koˈjusi	'his ear'
aúeu	aˈuju	'early'
aueúhe	auˈjuhe	'earlier'
híoa	ˈhiwa	'work'
eahioáa	eahiˈwaː	'to have work'
fúeŋu	ˈfujŋu	'hill'
epúeu	eˈpuju	'to fall'
céoa	ˈtʃewa	'earth'
eóu	ˈjou	'big fish'
éou	ˈewu	'thief'

(Szakos transcribes the last two examples with word initial glottal stops)

Taken together, the examples above show that certain midvowels are treated as syllabic, they receive stress and participate in morphophonemic processes such as reduplication, while at the same time, what have been transcribed by Tung as midvowels are treated as if they are non-syllabic as far as processes of stress and reduplication are concerned. These facts are what led Tung to conclude that Tsou had both syllabic and non-syllabic midvowels and led Li and Ho to propose the glides j and w. Tung's proposal is motivated by two factors. The first is that he finds it more parsimonious not

to add glides to the phonemic inventory. The second reason is that, at least in the case of j, following a consonant or following the front vowels i and e the glide sounds more like e (Li, 1979). The change in quality does not change the glide's inability to bear stress; thus, the variation in the glide's quality can be viewed as allophonic variation. Elsewhere, in word initial and word final and intervocalic position the auditory impression of both glides is similar to that of high vowels. From a phonological point of view, a glide is essentially a non syllabic vowel, and therefore in some sense Tung, Li and Ho are in agreement. The problem arises in Tung's transcription in which he fails to explicitly mark the difference between a syllabic and a non-syllabic vowel.

The wordlist used for making the recordings of consonants did not include examples of j and w in positions that are equivalent to the other consonants. However, several of the words in the wordlists do illustrate j word internally and in word initial position. Therefore, a preliminary acoustic study of j can be made. By chance w did not appear in word initial position with the exception of a single name. Therefore, it is not possible to perform a comparable analysis of w. However, as can be seen in Figure 14 above, its intervocalic characteristics are typical of a glide. Figure 17 below illustrates the highfront glide in word initial position and Figure 18 illustrates the allophonic midvowel characteristics of the front glide following a consonant.

When the front glide is in word initial or intervocalic position, it typically shows no steady state. Furthermore its F2 and F1 transitions begin at points that are typical of the high front vowel i. However, when the front glide is found in post consonantal position, the maximum frequency for F2 is lower and the minimum frequency for F1 is higher than in onset position. Altogether, the spectral characteristics for the front glide in post vocalic position place it nearer the midhigh front vowel e rather than i. Further study of the front glide is needed to determine the exact phonetic characteristics of the front glide. But our preliminary investigation suggests that it is similar to the English glide in 'yellow, a yellow ball, unite, reunite'. Since the convention is to transcribe a front glide as j in the absence of any height contrast, it is transcribed as such here.

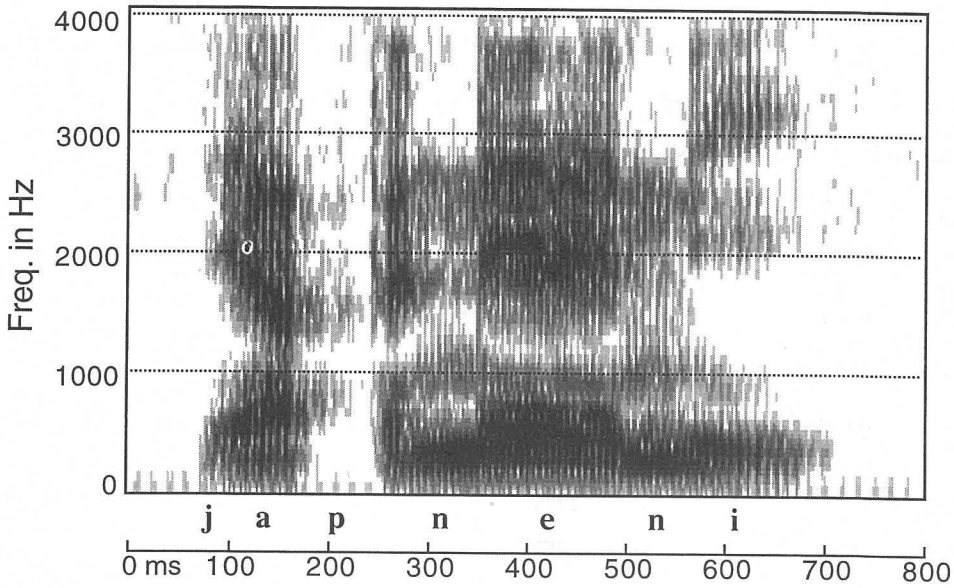


Figure 17: Example spectrogram showing j **in** japneni **'to sow'**

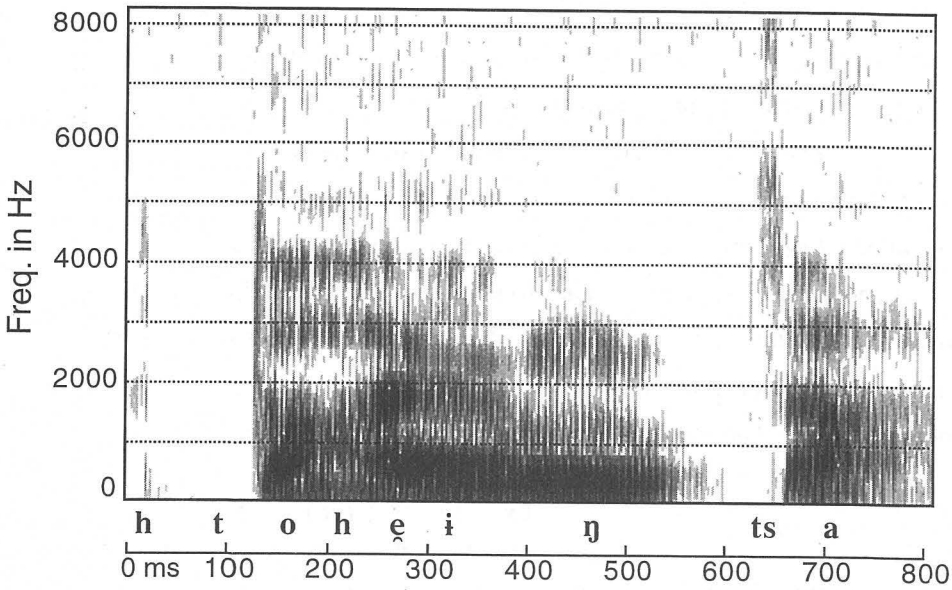


Figure 18: Spectrogram of htohjiŋsa [htohệŋtsa] **'stone to death' illustrating the midvowel characteristics of the front glide in post consonantal position.**

Stress

With the exception of a handful of words, stress is predictable in Tsou. As was seen in the examples above, stress tends to fall on the penultimate syllable.

ex: (From Tung)

óko	ˈoko	'child'
okósi	okˈosi	'his child'
mamespíŋi	mamesˈpiŋi	'woman'

This, however, is not the full story. When a word ends in a long vowel, stress falls on the last syllable.

ex: (From Tung)

emóo	eˈmoː	'house'
eémo	ˈeːmo	'to build a house'

The full generalization, then, is that stress falls on the final syllable if it has a long vowel, else on the penultimate syllable. This pattern is one of Tung's motivations for describing long vowels as disyllabic; he wanted to say that stress was always on the penultimate syllable. Tung, however, was writing before the advent of metrical theories of stress such as that described in detail in Hayes (1993) and thus lacked the tools for analyzing this pattern as moraic stress. In a theory of moraic stress, a syllable containing a long vowel is heavy and bears two morae. A syllable with a single vowel is light and bears a single mora. The stress pattern in Tsou is characteristic of moraic trochees: main stress falls on the final syllable if it is heavy, else on the penultimate syllable. Because the main stress falls at the right edge of a word, stress is assigned from right to left. The moraic trochee foot template is as follows:

(x	.)	(x)	
∪	∪	or	–

Where ∪ is a light syllable (μ) and – is a heavy syllable (μμ)

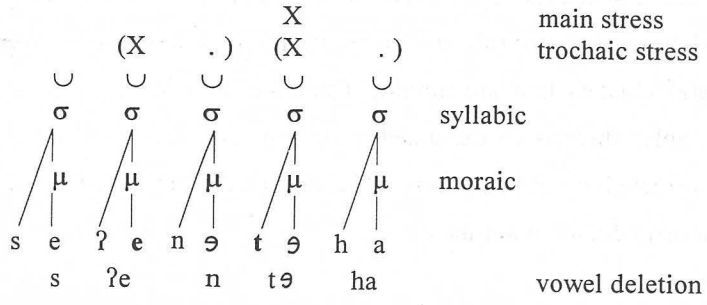
ex:

(x .) óko	'child'
(x .) okósi	'his child'
(x) emóo	'house'
(x) eémo	'to build a house'

Evidence for moraic trochees comes from the Luhtu dialect of Tsou which has a productive process of vowel deletion described in Tsuchida. Alternating vowels preceding the main stress are deleted. Vowel deletion does not occur if a cluster of three or more consonants would occur, or if the vowel is in an onsetless word initial syllable or in word final position. Vowel deletion also fails to occur if the vowel in question is long, i.e. if the syllable is heavy. This pattern is also characteristic of moraic trochees. Every heavy syllable receives stress, every other light syllable receives stress, all stressless syllables excluding those mentioned above are deleted. Our data were collected without this point in mind, and do not allow us to fully determine vowel deletion rules in the Tfuea dialect. From what we have observed it is clear that vowel deletion is an active process in Tfuea, and it seems likely that both dialects have the same stress pattern: main stress is on a final heavy syllable else on the penultimate syllable.

ex: from Tsuchida

se?e- nɘtɘh- a [s?en'tɘha] 'to cut off with a bolo' (general focus)



Richard Wright & Peter Ladefoged

Consonant clusters

One of the most interesting aspects of Tsou is the wide array of consonant clusters that are permissible in word initial position. In general a consonant cluster is made up of two and only two consonants, and almost any two of the Tsou single consonants may appear in either order. There are, however, a few exceptions to this generalization. With the exception of h and ʔ, homorganic fricative-stop or stop-fricative clusters do not occur. If h is regarded as not having a glottal place of articulation, hʔ and ʔh need not be considered exceptions. Strengthening the notion that in the phonological definition of possible clusters h should be regarded as a velar fricative is the fact that ph and th do occur as initial and medial clusters, but kh occurs only in Japanese loans (but see Szakos, 1994 and Wright, 1996). Of the consonants with oral constrictions only nasals precede homorganic stops or fricatives. The preglottalized alveolar implosive ɗ clusters only with the glottal stop. As most of the data was compiled using Tung's text where the glides are unmarked, no systematic data was collected on the clustering patterns of j or w. Not all of the consonant clusters that are possible in Tsou have been documented in word initial position. Tung's description is corpus based, and therefore any absence of a particular cluster may very well be due to chance. The present study expands somewhat on Tung's list of initial consonant clusters, yet any absences may too be due to chance. Some time was spent trying to elicit words that begin with the clusters that are listed but not exemplified in Tung's text. When this process failed it is an indication that the cluster is at least dispreferred, but it does not rule out the possibility of such a cluster. Szakos (1994) includes several clusters that are missing from the data in this study. The Appendix contains a table which is an expanded version of that found in Tung. The glosses have been corrected where necessary. The following table illustrates the combinations of consonants found in our data.

C1 ↓ \ C2 →	p	f	v	ɓ	m	t	ɗ	ts	s	z	n	k	ŋ	ʔ	h
p						x		x	x	(x)	x	x	x	x	x
f						x		x			x	x	x	x	
v								x			(x)			(x)	x
ɓ											x	(x)			
m	x	x		(x)		(x)		x	x	x	x			x	x
t	x	x	x	(x)	x						x	x	x	x	x
ɗ															
ts	x	x	x	(x)	x						x	x	x	x	x
s	x		x	x	x						x	x	x	x	
z														(x)	
n	(x)				x	x		(x)	x	(x)					(x)
k								(x)	x		x			(x)	
ŋ			x		(x)	(x)		(x)	(x)	(x)		(x)			x
ʔ	x	(x)	x		x	x		x	x		(x)	(x)			(x)
h	x		x		x	x		x	(x)	(x)	x	(x)	(x)	x	

Table 5: Consonant clusters attested in the data
x = clusters that appear in word initial and word internal position
(x) = clusters that appear only in word internal position

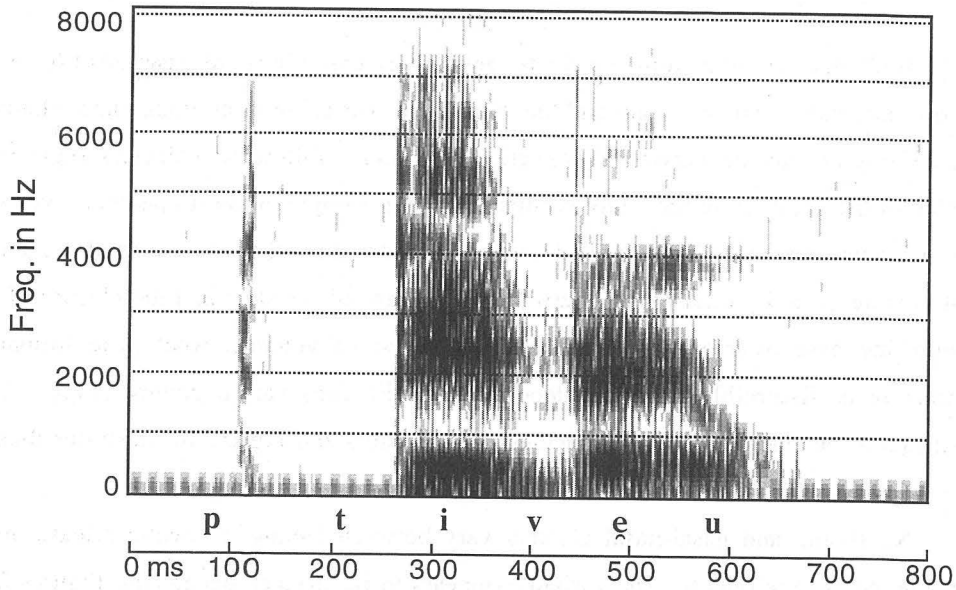


Figure 19: Example spectrogram showing the characteristic voiceless release of voiceless stop clusters in ptiveu̯ 'a kind of reed'

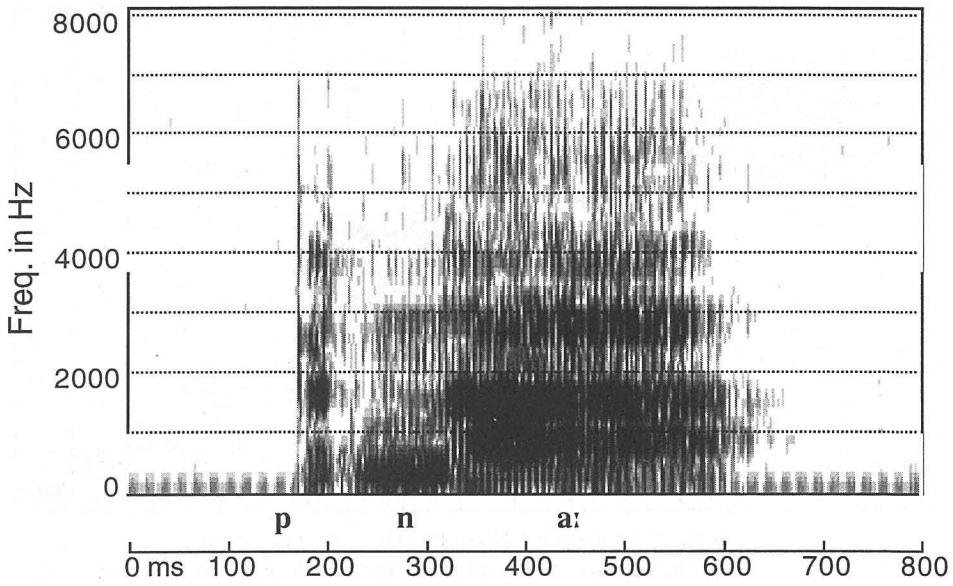


Figure 20: Example spectrogram showing the characteristic vocalic release in voiceless stop + nasal clusters in pna: **'to shoot'**

Both members of a stop-stop cluster are fully released in word onset position and word internally. There is an audible burst that varies in amplitude, and almost invariably no voicing between a voiceless stop and a following voiceless segment. (One of the speakers in the study produced voicing forming a short epenthetic vowel in this environment). When a voiceless stop precedes a nasal there is a small amount of voicing, 3 to 5 glottal pulses, between the release of the stop and the closure of a following nasal segment. During the brief period of voicing, what little formant structure is discernible is highly dependent on the consonantal context. Figure 19 illustrates the characteristics of voiceless stop clusters and Figure 20 illustrates those of voiceless stop plus nasal clusters.

Nasal-stop and nasal-nasal clusters vary between having a vocalic release and having no vocalic release. The variation appears to be speaker dependent. Figures 20 and 21 illustrate the two possibilities.

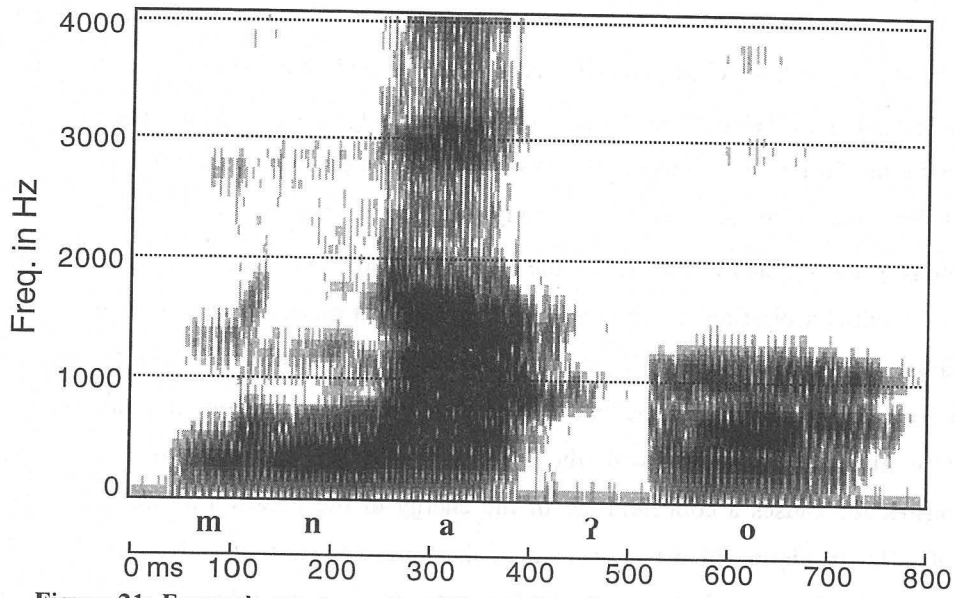


Figure 21: Example spectrogram with vocalic release in a nasal-nasal cluster in mnaʔo **'come to do one thing'**

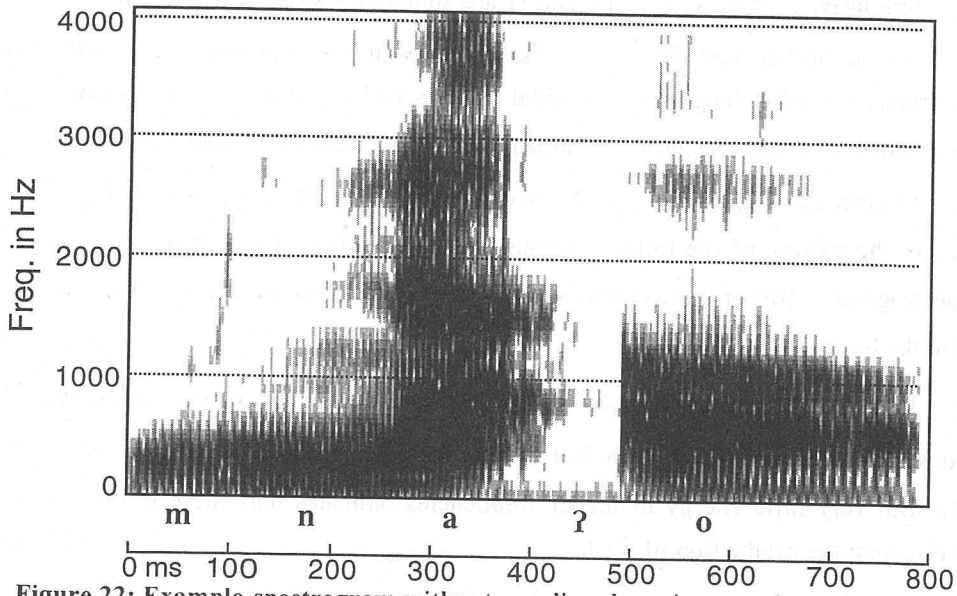


Figure 22: Example spectrogram without vocalic release in a nasal-nasal cluster in mnaʔo **'come to do one thing'**

Some of the most unusual clusters involve h and ʔ. Clusters such as ph and th, (but not, as we have observed, kh) occur, and are much like aspirated consonants in other languages. What is, to the best of our knowledge, unknown in other languages are initial clusters with h as the first member, such as ht, illustrated in Figures 23 and 24. For most of the speakers in the study the frication associated with the h has low intensity and spreads widely across the spectrum, as illustrated in Figure 23. In Figure 24 the initial aspiration is somewhat stronger than that produced by the other speakers (chosen so that the acoustic characteristics were clearly visible).

Before the alveolar stop there is a movement so that the acoustic structure becomes more like that associated with the alveolar burst: the narrowing of the alveolar constriction causes a concentration of the energy of the frication between 2000 and 5000 Hz. In Figure 24 the movement of the tongue towards the alveolar closure is clear in the change in frequency of the formants within the frication.

Similarly, clusters such as pʔ and tʔ are sometimes sequences of a plosive and a glottal stop, but are sometimes much like ejectives in other languages; and again what is unknown in other languages are initial clusters such as ʔp and ʔt, illustrated with ʔt in Figure 25. Perhaps the most exceptional of all is the cluster hʔ illustrated in Figure 26.

In clusters of glottal stop and voiceless stop the glottal stop is usually released before the closure of the following consonant is complete. Often this is visible on a spectrogram as two or three creaky glottal pulses. Occasionally the creakiness persists into the beginning of the vowel.

In the hʔ cluster, the formants visible in the frication show coarticulation with the vowel following the glottal stop. In this case the concentration of energy around 1000 Hz with very little energy in higher frequencies indicates that there is lip rounding throughout the production of the h.

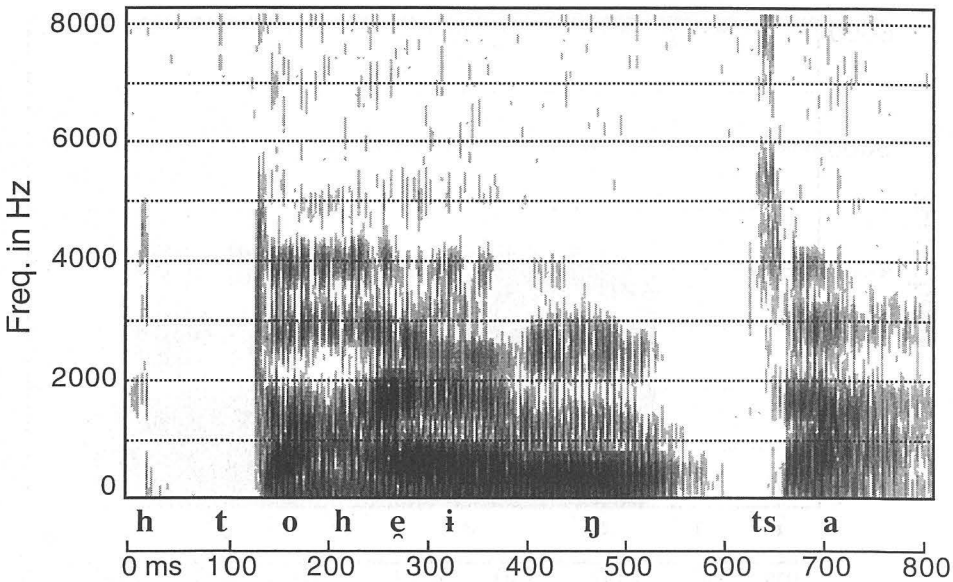


Figure 23: Spectrogram of htohjɨŋtsa [htohệŋtsa] **'stone to death' illustrating the typical low intensity, broad spectrum frication with concentration of energy at the closure of the following stop.**

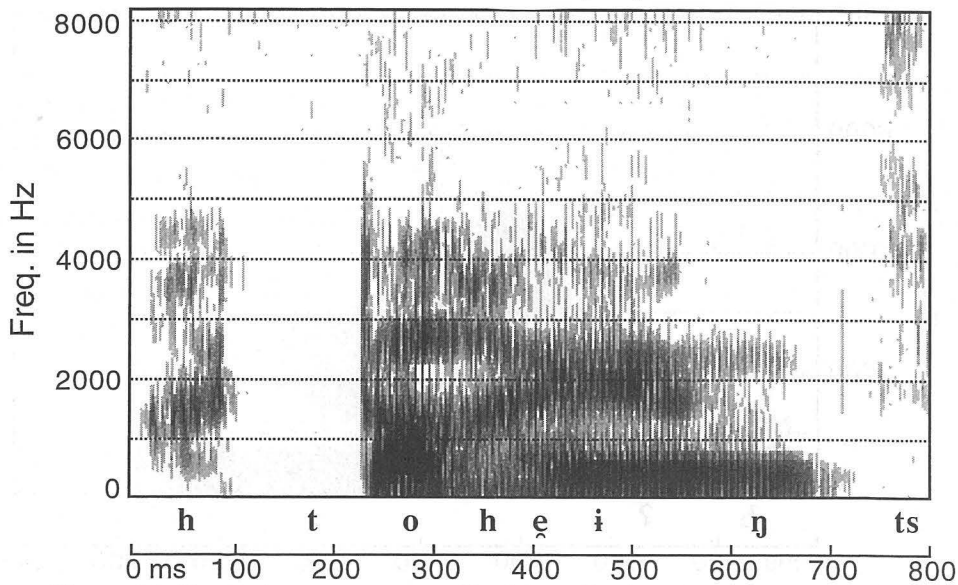


Figure 24: Example spectrogram of htohjɨŋsa [htohệŋtsa] **'stone to death' illustrating stronger frication.**

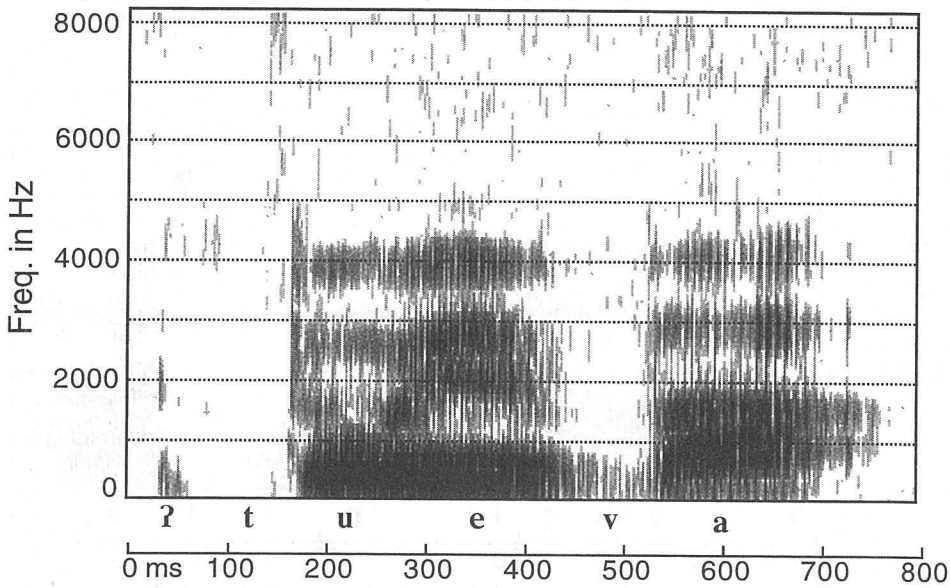


Figure 25: Example spectrogram of ʔtueva 'the third month' illustrating the characteristic glottal release .

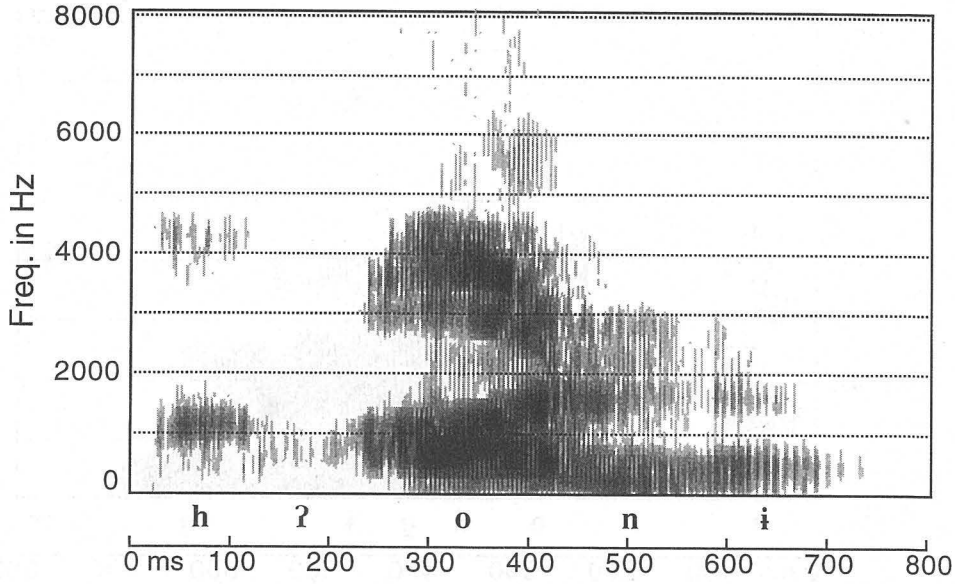


Figure 26: Example spectrogram of hʔoni̯ 'liver' showing h coarticulation with the vowel through the glottal stop.

Conclusion

This paper has outlined the main phonetic characteristics of Tsou. But there are many points of phonetic interest still to be investigated. In particular, further research on the consonant clusters is needed. We are not completely certain whether there are some consonant combinations that are possible but do not appear in our data. In previous analyses, all syllables are said to have the form (C)(C)V so that there are no codas in the language, regardless of the position of the syllable within the word. These analyses are based in large part on the fact that there are no consonants at the ends of words. Our data show that in medial position the first voiceless plosive in a cluster is released which is a characteristic normally associated with onsets. We need to conduct further investigations to establish the full range of consonant clusters in Tsou. In addition, further investigation of the nature of unusual clusters like those in Tsou will shed light on questions of the nature of the makeup of possible syllable onset clusters and the role of saliency in determining ordering of articulations in speech.

(Accepted for publication 27 February 1997)

References

Fant, Gunnar (1973). *Speech sounds and features*. Cambridge, MA, MIT Press.

Fischer-Jørgensen, Eli (1954). Acoustic analysis of stop consonants. *Miscellanea Phonetica* **2**: 42-59.

Fuller, Michael (1990). Pulmonic ingressive fricatives in Tsou. *Journal of the International Phonetic Association* **20**: 9-14.

Hayes, Bruce (1995). *Metrical Stress Theory: Principles and case studies*., University of Chicago Press, Chicago.

Ho, Dah-an (1976). Tsou phonology. *Bulletin of the Institute of History and Philology, Academia Sinica* **47**: 245-74.

Ladefoged, Peter (1993). Linguistic phonetic fieldwork: a practical guide. *UCLA Working Papers in Phonetics* **84**: 1-24.

Ladefoged, Peter and Zeitoun, Elizabeth (1993). Pulmonic ingressives phones do not occur in Tsou. *Journal of the International Phonetic Association* **23**(1): 13-15.

Li, Paul J-K. (1979). Variations in the Tsou dialects. *Bulletin of the Institute of History and Philology, Academia Sinica*. **50**: 273-96.

Lin, Hen-li (1955). On the quasi indirect passive voice in the Tsou grammar in Ali Mountain. *Journal of Chinese Ethnography* **1**: 207-216.

Lindau, Mona (1984). Phonetic differences in glottalic consonants. *Journal of Phonetics* **12**: 147-155.

McCarthy, John J. & Prince, Alan S. (ms). *Prosodic Morphology: Constraint interaction and satisfaction*. University of Massachusetts, Amherst and Rutgers University.

Nevskij, N.A. (1935). *Materialy po Govoram Jazyka Cou*. Moscow, Trudy Instituta Vostokovedenija.

Ogawa, Naoyoshi & Asai, Erin (1935). *The myths and traditions of the Formosal native tribes*. Taipei, Taihoku.

Starosta, Stanley (1974). Causative verbs in Formosan languages. *Oceanic Linguistics* **13**: 279-369.

Szakos, Joseph (1994). *Die Sprache der Cou: Untersuchungen zur Synchronie einer austronesischen Sprache auf Taiwan*. Ph. D. Thesis, Rheinischen Friedrich-Wilhelms-Universität, Bonn.

Tsuchida, Shigeru (1972). The origins of the Tsou phonemes /b/ and /d/. *Gengo Kenkyu* **46**: 42-52.

Tsuchida, Shigeru (1976). *Reconstruction of Proto-Tsouic phonology*. Tokyo, Studies of languages and culture of Asia and Africa, Monograph Series.

Tung, T'ung-ho (1964). *A Descriptive Study of the Tsou Language*. Taipei: Institute of History and Philology, Academia Sinica Special Publication.

Wei , Hwei-lin, Yu, Chin-chuan & Li, Hen-li (1952). *Ethnography of the Ethnic Group Tsou*. Taipei: Commission of Historical Research of Taiwan.

Wright, Richard (1996). *Consonant clusters and cue preservation in Tsou*. Ph. D. thesis, UCLA.

Appendix

I: Word List used in Vowel Recordings

IPA	GLOSS	IPA	GLOSS
pakʔi	'bad, wicked'	paʔti	'to show, to let see'
peːɗa	'can, be able to'	pepe	'sky, heaven'
pitu	'seven'	pio	'how many'
pohe	'maize'	poʔe	'to pray making wine'
puzu	'fire'	puːtu	'Chinese'
taso	'strong'	tafʔu	'kind of skirt'
teːɗu	'to arrive on time'	teːʃi	'rope, cord'
timʒi	'to hold'	titha	'to use'
tofɨ	'gourd'	toti	'kind of fungus'
tufku	'to wash clothes'	tutu	'seed, grain, ball'
tɨsʔɨ	'clear (of water)'	tɨʔsɨ	'arrow'

II: Word List used in Consonant Recordings

paʔmiʒi	'to let pass'	peʔpiʔi	'eyebrow'
baʔi	'grandmother'	bjahtʃi	'fruit, flesh'
maʔkuvʔo	'to surprise'	majo	'to take'
meoiʃi	'large, big'		
faʔe	'a man's name'	fjaŋo	'body'
vaʔhɨ	'stream'	vjovjo	'kind of deer'
taʔtsɨ	'lamb'	teʔi	'to shit'
ɗauja	'maple'	ɗiŋki	'mud'
navju	'cooked rice'	neʔe	'there (out of sight)'
tsaʔi	'dirty'	tʃei	'dream'

saptʃi	'face'	ʃeu	'intestine'
zomɨ	'bird'		
kadĩ	'knee'	keːi	'eye secretion'
ŋajo	'mouth'	ŋeːi	'hemp'
hahʔo	'friends'		

III: CONSONANT CLUSTERS IN THE PRESENT DATA

CLUSTER	INITIAL	GLOSS	MEDIAL	GLOSS
pt	ptiveu	'a kind of reed'	sɨptɨ	'four'
pk	pkaːko	'to escape'	------	-------
pʔ	pʔuŋu	'body joint (knee)'	sopʔo	'light (weight)'
pn	pnaː	'to shoot'	japneni	'to sow'
pŋ	pŋiji	'pestle'	apŋu	'a basket'
pts	ptsoknɨ	'tall bamboo'	optsoi	'to kill'
ps	psipŋa	'to thrust into'	jopʃi	'dusk'
pz	------	-------	ŋipza	'mildew'
ph	phiŋi	'door'	tsaphɨ	'palm, sole, paw'

t

tp	tposɨ	'drawing'	atpɨta	'to win'
tɓ	------	-------	ketbɨ	'storage'
tk	tkiti	'miser'	kitkiti	'wrist'
tʔ	tʔaŋo	'foot'	atʔiŋhi	'but'
tm	tmalɨ	'to hear'	matmomane	'sing for a long time'
tn	tnɨjɨ	'rain'	nɨtnɨ	'lung'
tŋ	tŋoː	'pot'	botŋonɨ	'many'
tf	tfuja	'place name'	fitfi	'edge'
tv	tvoʔhusa	'bamboo tube'	teakotva	'narrow'
th	thoepza	'to elevate'	aothomɨ	'to try'

k

k?			ak?i	'grandfather'
kn	knuju	'to cheat'	konakn̈	'species of plant'
kts	------	-------	zotɨkc̈	'to press'
ks	ksiksi	'burning charcoal'	ksiksi	'burning charcoal'

?

?p	?pitva	'monkey'	pe?pi?i	'eyebrow'
?t	?tujva	'the 3rd month'	aopa?to	'to appear'
?d̈	------	-------	ka?d̈i	'knee'
?k	------	-------	ko?ko	'therefore'
?m	?maska	'the 10th month'	ma?mio	'to think'
?n			na?no	'very'
?ts	?tsonza	'the 1st month'	smuju?tsu	'to pierce'
?f			tmau?fu	'to stir'
?v	?voja	'the 8th month'	i?vaho	'again'
?s	?sia	'the 9th month'	to?so	'to throw'
?h	------	-------	va?hɨ	'small river'
ɓk	------	-------	tiɓkoɓkotsa	'therefore'

m

mp	mpɨtɨ	'to hold'		
mɓ	------	-------	se?imɓa	'five hundred'
mt	------	-------	aɨmtɨ	'actually'
m?	m?um?u	'body hair'	am?uhu	'early, quickly'
mn	mna?o	'come to do one thing'	ɨmnɨ	'good'

mts	mtsoː	'eye'	mamtsino	'to bathe'
mf	mfeɨʔsɨ	'to cover'		
ms	msapie	'put on shoe'	nomso	'have enough room'
mz	mzaː	'our'	omza	'upper side'
mh	mhino	'to buy'	mimho	'to permit'

n

np	------	-------	aɨnpɨnpɨ	'at will'
nt	nteʔo	'1 sg +conditional prefix'	tʔuantɨhɨ	'star's name'
nk	------	-------	kunkunu	'wooden water jar'
nʔ	------	-------	tanʔe	'here'
nm	nmɨjo	'adult animal'	anma	'sixth month'
nts	------	-------	maintsa	'to say (agent focus)'
nv	------	-------	sinvi	'to chew'
ns	nsoː	'pond'	zonso	'lie in wait for an animal'
nz	------	-------	ʔtsonza	'first month'
nh	------	-------	inhe	'sow'

ŋ

ŋt	------	-------	hoŋtuŋtu	'to burn one end'
ŋk	------	-------	oŋko	'name'
ŋʔ	------	-------	aŋʔosɨ	'gang up on'
ŋm	------	-------	joŋma	'to be in ambush'
ŋts	------	-------	jɨŋtsa	'mad'
ŋv	ŋvoe	'to dry in the sun'	tuŋva	'to powder'
ŋs	------	-------	moŋsi	'to cry'
ŋz	------	-------	meŋzu	'lance'
ŋh	ŋhou	'monkey'	atʔiŋhi	'only'

ts

tsp	tspɨʔho	'hip'	tsputspuhu	'arrival in the mountains'
tsb	------	-------	etsbɨhɨ	'to feel heavy'
tsk	tskosi	'to level'	putsku	'navel'
tsʔ	tsʔojha	'stream'	jatsʔɨ	'to stand'
tsm	tsmoi	'bear'	hutsma	'day before, day after'
tsn	tsnɨmɨ	'banana'	a:tsni	'surely'
tsŋ	tsŋuhu	'chestnut'	atsŋihi	'continuously'
tsf	tsfuo	'stomach'	teotsfuŋa	'convergence of streams'
tsv	tsvoʔha	'pierce'	tatsvohʔi	'long'
tsh	tshumu	'water'	botshio	'to know'

f

ft	ftuke	'to slouch'	mofti?i	'to flash'
fk	fkoi	'snake'	tsofkoja	'clean'
fʔ	fʔue	'sweet potato'	emufʔo	'to bury'
fn	------	-------	jofna	'evening'
fŋ	fŋú:	'head'	emɨfŋɨ	'to hide'
fts	ftsuju	'egg'	faftsuja	'a castrated man'

v

vn	------	-------	ɓɨvnɨ	'flower'
vʔ	------	-------	avʔu	'dog'
vts	vtsoŋɨ	'spouse'	tsivtsi	'tail'
vh	vhonɨ	'lean'	tsovhi	'distant'

s

sp	spoja	'blanket'	mamespiŋi	'woman'
sɓ	sɓuku	'bamboo shoot'	------	-------
sk	skoʔa	'to hug while sleeping'	joskɨ	'fish'
sʔ	sʔofɨ	'stick'	kosʔoza	'shrimp'
sm	smu:	'dew'	esmi	'to pass by'
sn	snɨfɨ	'skin'	esnɨthɨ	'to cut and break'
sŋ	sŋɨsŋɨjo	'a place name'	asŋɨtsɨ	'constantly'
sv	svɨjɨ	'earring'	majasvi	'dance'

z

zʔ	------	-------	mahizʔo	'poisoning'

h

hp	hpɨhpɨŋɨ	'world'	sohpoi	'distressed'
ht	htoeuŋsa	'repeated'	ehti	'branch'
hk			kuhku	'fox'
hʔ	hʔonɨ	'liver'	hʔohʔo	'a wound'
hm	hmuju	'blood'	tsohmo	'to pour in'
hn	hnou	'small house'	smuhnu	'to dispatch'
hŋ	------	-------	oahŋu	'relatives'
hts	htsuju	'hill'	bjahtsi	'fruit'
hv	hvosɨ	'enlarged genitals'	etohva	'smoking pipe'
hs	------	-------	ɓohsopɨ	'to stick in'
hz	------	-------	tsohzona	'noon'

鄒語語音研究

Richard Wright(李昌豪)/ Peter Ladefoged

印地安那大學心理學系/ 加州大學洛杉磯分校語言學系

本文在研究鄒語特富野方言的語音特徵。作者兼採聲學和聽學的分析，因此是首次以儀器研究鄒語的著作。因爲詳細探討鄒語的各種輔音和元音，所以過去學者對於鄒語語音描述的歧異之處，在本文都獲得澄清。本論文分析包括元音共震峰的測量、輔音的帶音、發聲部位的嗓音起始時間、元音的基本音高。除了音段的描述以外，對於輔音群也有初步的探討。鄒語輔音群在世界上的語言是很少見的。

關鍵詞：鄒語 聲學語音學 輔音數 元音數 輔音群