Ersu

Katia Chirkova

Centre de Recherches Linguistiques sur l'Asie Orientale,
Centre National de la Recherche Scientifique, Paris, France
katia.chirkova@gmail.com

Dehe Wang

Xichang College, Xichang, China
365872711@qq.com

Yiya Chen

Leiden University Centre for Linguistics,
Leiden Institute for Brain and Cognition, Leiden, The Netherlands
yiya.chen@hum.leidenuniv.nl

Angélique Amelot

Laboratory of Phonology and Phonetics,
Centre National de la Recherche Scientifique, Paris, France
angelique.amelot@gmail.com

Tanja Kocjančič Antolík

Laboratory of Phonology and Phonetics,
Centre National de la Recherche Scientifique, Paris, France
tkocjancic@googlemail.com

The Ersu language (/ɔ̂-ʂý xò/, 尔苏语 *ěrsūyǔ*, ISO-639 code *ers*) is spoken by approximately 16,800 people who reside in five counties in Sichuan Province (四川省) in the People's Republic of China: (i) Ganluo (甘洛县), and (ii) Yuexi (越西县) counties of Liangshan Yi Autonomous Prefecture (凉山彝族自治州), (iii) Shimian (石棉县) and (iv) Hanyuan (汉源县) counties of Ya'an Municipality (雅安市), and (iv) Jiulong (九龙县, Written Tibetan, hereafter WT *brgyad zur*) county of Ganzi (甘孜, WT *dkar mdzes*) Tibetan Autonomous Prefecture.[1] Ersu has two closely related sister languages: Lizu (/li⁵⁵-zu⁵⁵-hũ⁵⁵/ or /ly⁵⁵-zu⁵⁵-hũ⁵⁵/, 里汝语 *lǐrǔyǔ* or 栗苏语 *lìsūyǔ*) and Duoxu (/do³³-ɕu³³-na³¹/, 多续语 *duōxùyǔ*

---

[1] In transcriptions of Ersu words '-' stands for morpheme boundary and ' = ' stands for clitic boundary. In transcriptions, the Low tone is marked as 'à', and the High tone is marked as 'á' (for more detail, see section 'Prosodic organisation' below).

or 多须语 *duōxūyǔ*).[2] Lizu is spoken in the neighbouring counties of (i) Muli Tibetan Autonomous County (木里藏族自治县, WT *smi li rang skyong rdzong*), (ii) Mianning (冕宁县), and (iii) Jiulong; whereas Duoxu is spoken in the county of Mianning, all in Sichuan province (see the map in Figure 1).
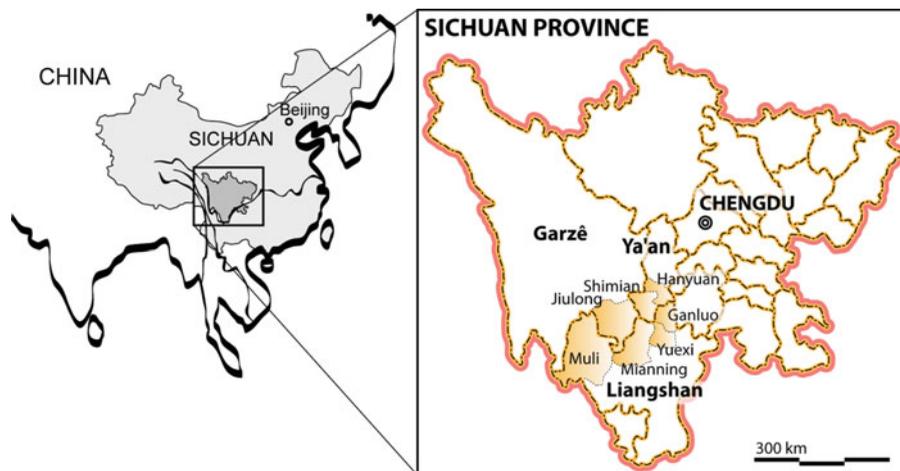

Figure 1 (Colour online) Distribution of the Ersu, Lizu, and Duoxu languages.

In present classifications of Tibeto-Burman languages spoken in Southwest China, Ersu, Lizu, and Duoxu are viewed as three dialects of one Ersu language. The Ersu language is, in turn, classified as a member of the Qiangic subgroup of the Tibeto-Burman language family (for more details, see Bradley 1997: 36–37; H. Sun 2001; Chirkova 2012). In this conception, Lizu is the western dialect of the Ersu language, Duoxu is its central dialect, and Ersu proper is the eastern dialect of the Ersu language. In contrast to this received view and in accordance with the fact that differences between Lizu, Duoxu, and Ersu surpass the limit of mutual intelligibility (H. Sun 1982, Chirkova 2014), we consider Lizu, Duoxu, and Ersu as separate languages, and not as dialects of one Ersu language (Yu 2012: 1). The phonological, lexical, and morphosyntactic differences between Ersu, Lizu, and Duoxu are likely to be in part due to the competing influences of the languages with which they are in contact. More specifically, Ersu has been historically influenced by Pumi (普米), Nuosu (Northern Ngwi or Yi 彝), and Mandarin Chinese (the local variety of Southwest Mandarin, hereafter SW Mandarin) (Wu Da 2010: 3). By contrast, Lizu has been influenced by Tibetan, Pumi, and Namuzi (纳木兹) languages (Chirkova & Chen 2013). Finally, Duoxu has been essentially influenced by SW Mandarin as well as by Nuosu (Chirkova 2014).

The main focus of this illustration is on a synchronic analysis of Ersu proper. (For more comparatively- and diachronically-oriented studies of Ersu, Lizu, and Duoxu, the interested reader is referred to Yu 2012, Chirkova & Handel 2013a, Chirkova 2014.)

Ersu is relatively little researched, but the group and its language and culture have been receiving increasing attention in recent years (e.g. Wang 2010, Wu Da 2010, Schmidt 2011, Zhang 2013, 2014). Early linguistic accounts include H. Sun (1982, 1983) and Liu (2007 [1983]), which focus on the Ersu as spoken in Ganluo county.

---

[2] Tone notation in Lizu is provided in superscript letters to the left of the lexical word. The following abbreviations are used: F = falling tone, R = rising tone, EP = equally-prominent pattern, LP = left-prominent pattern, RP = right-prominent pattern (see Chirkova & Chen 2013 for a detailed discussion).

Ersu is an endangered language. It is essentially used as the primary language of oral communication in family and community events. Older Ersu speakers (typically above their sixties) are mostly trilingual (Ersu, SW Mandarin, Nuosu). Over the last three decades, most Ersu speakers have been bilingual using SW Mandarin in daily life. The current trend for the school-going generation is to become practically monolingual in Mandarin. Ersu has its own pictographic writing system, known as *shaba* 沙巴 (Ersu /ṣàpá/'ritual priest') writing, which is chiefly used by Bon priests (e.g. H. Sun 2009, Wang 2011).

The present illustration provides a preliminary description of Ersu on the basis of data from three speakers: two male speakers in their early sixties and one female speaker in her early forties, all born and raised in Ganluo county (Zela Township 则拉, Liangshan group 凉山组, Mofanggou village 磨坊沟村 /dʐéɬá lò, ẓákáɻʐgv́ fv́, nóNdʐv́ pá/). Given the phonetic complexity of the consonant and vowel sounds of Ersu (including a number of typologically uncommon trilled retroflex sounds and phonemic fricative vowels), further research, based on more speakers, is required for a comprehensive analysis of this language. In the present Illustration, the basic phonetic characteristics of Ersu are described through acoustic, palatographic, aerodynamic, electroglottographic, and video data. We have chosen to illustrate the discussion with audio files (the word list and the text provided in the present paper) as read by the second author, a male native speaker of Ersu. This is because we were in the fortunate position of recording him in a phonetics laboratory, yielding high quality audio recordings. Conversely, palatographic images in the text are from the female speaker. This is because, as the youngest speaker among our language consultants, her dental and palatal condition yielded the clearest images.

Consonants

	Bilabial	Labio-dental	Dental	Alveolar	Alveolo-palatal	Retroflex	Velar
Plosive	p pʰ b		t tʰ d				k kʰ g
Affricate			t̪s̪ t̪s̪ʰ d̪z̪	tʂ tʂʰ dʐ	tɕ tɕʰ dʑ	tɽ tɽʰ dɽ	
Nasal	m		n				ŋ
Trill						ɽ	
Fricative		f v	s̪ z̪	ʂ ʐ	ɕ ʑ		x
Approximant	w				j		
Lateral				l			
Lateral fricative				ɬ			

The Ersu consonant inventory consists of 38 phonemes. There is a general three-way manner distinction in stops and affricates: voiceless unaspirated, voiceless aspirated, and voiced. Ersu has an extensive system of coronal consonant contrasts in affricates at the dental, alveolar, alveolopalatal, and retroflex places of articulation. Alveolopalatals are marginal. They mainly occur in (recent) loanwords from Mandarin Chinese, such as /kʰà-tɕà-tɕá/ 'to pick up with chopsticks' (Chinese *jiā* 夹, SW Mandarin /tɕia⁴⁴/), /kòtɕó/ 'legging, puttee' (Chinese *guǒjiǎo* 裹脚, SW Mandarin /ko⁵³tɕio²¹/), /ɕá/ 'incense' (Chinese *xiāng* 香, SW Mandarin /ɕian⁴⁴/).[3] In the native vocabulary, alveolopalatals have a restricted distribution, co-occurring only with the vowels /i o a/. Examples include /tɕí-tɕí/ 'to squeeze, to pick up with chopsticks; to cut with scissors' (possibly, a loanword from Chinese *jiǎn* 剪 'to cut with scissors', SW Mandarin /tɕian⁵³/), /tɕó/ 'to wrap, to bind', /ódʑá/ 'pear' (see below on vowels).

---

[3] SW Mandarin readings are based on Li (2010) and provided in the original transcription.

p	pɛ́	'to give as a present'	ʂ	ʂɛ́	'louse'
pʰ	pʰɛ́	'male'	ʐ	ʐɛ́	'to flow'
b	bɛ́	'insect'	l	lá	'fertilizer, manure'
m	mɛ́	'ink'	ɬ	ɬá	'month'
w	wá	'Chinese yam'	tɕ	tɕó	'to wrap, to bind'
f	fí	'mildew'	tɕʰ	tɕʰó	'pepper'
v	ví	'tinder'	dʑ	dʑó	'to push'
t	tɛ́	'one'	ɕ	ɕó	'to sweep'
tʰ	tʰɛ́	'he, she, it'	ʑ	ʑó	'to soften (skin)'
d	dɛ́	'to weave'	j	jó	'sheep'
t̪s̪	t̪s̪ɛ́	'cloud'	tʈ	tʈɛ́	'steelyard'
t̪s̪ʰ	t̪s̪ʰɛ́	'to wash'	tʈʰ	tʈʰɛ́	'fence'
d̪z̪	d̪z̪ɛ́	'rooster (zodiac sign)'	dʈ	dʈɛ́	'pair'
n	nɛ́	'you, thou'	ʈ	ʈɛ́	'earth, soil'
s̪	s̪ɛ́	'air'	k	ká	'to be stupid'
z̪	z̪è	'wife'	kʰ	kʰá	'barley'
ʦ	ʦó	'oats'	g	gá	'to sing'
ʦʰ	ʦʰó	'to rot (of wood)'	ŋ	ŋá	'to be hungry'
ʣ	ʣó	'water'	x	xá	'to have'

We studied the four-way contrast in coronals using palatographic analysis techniques (as described in Marchal 1988, Ladefoged 2003: 36–42, and Anderson 2008). The list of words used in the palatographic study consisted of monosyllabic words in common use in the language, each of which included only one coronal consonant. The palatography procedures were to paint the tip, blade, and front of the tongue with a solution of one part olive oil and one part of finely ground activated charcoal. After each word was pronounced, a mirror was placed in the speaker's mouth, resting against the lower teeth and the reflection of the upper palate and teeth was recorded using a video camera.

Stops and affricates

Ersu dental stops and affricates are both produced in the dental region. Dental stops involve contact on both the teeth and most of the alveolar ridge, making them (laminal) denti-alveolar. By contrast, dental affricates involve a smaller contact area, which includes the upper front teeth and the front part of the alveolar ridge. This is illustrated in Figure 2, with the words /dɛ́/ 'to weave' and /d̪z̪ɛ́/ 'rooster (zodiac sign)'.

Alveolar affricates are produced with the tongue touching the middle of the alveolar ridge (laminal flat alveolar). The contrast between dental and alveolar affricates is illustrated in Figure 3, with the minimal pair /d̪z̪ó/ 'to lay bricks' and /ʣó/ 'water'.

Alveolopalatal affricates are produced with the blade of the tongue behind the alveolar ridge and the body of the tongue raised towards the palate, thus involving simultaneous alveolar and palatal articulation (compare 'laminal palatalized post-alveolar', as described in Ladefoged & Maddieson 1996: 153–154). The contrast between alveolar and alveolopalatal affricates is illustrated in Figure 4, with the words /ʣó/ 'water' and /dʑó/ 'to push'.

Ersu retroflex affricates are produced with the point of contact on the roof of the mouth, that is, in the hard palate. The contact is made with the underside of the tongue (subapical). The articulation of retroflex affricates involves lateral bracing of the tongue against the teeth, so that the tongue tip is free to move to and from the hard palate. Figure 5 contains palatograms of the words /tʈò/ 'gallbladder' and /dʈò/ 'pot, pan'. As we did not paint the underside of the tongue with a mixture of olive oil and activated charcoal powder, the area of contact is not visible on the image.

Ersu retroflex affricates have a trill release. They involve an aerodynamically induced movement of the tip of the tongue, causing intermittent contact between the tip of the tongue

/dɛ́/ 'to weave'   /d̠ʐɛ́/ 'rooster (zodiac sign)'
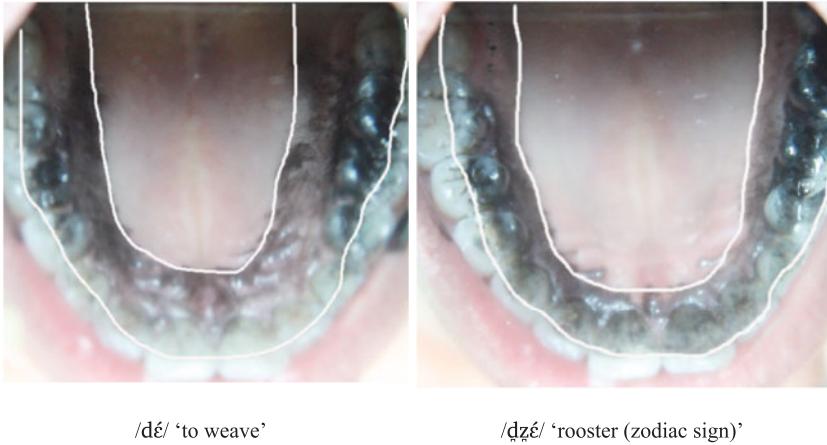
Figure 2 (Colour online) Palatograms of the words /dɛ́/ 'to weave' and /d̠ʐɛ́/ 'rooster (zodiac sign)'.

/d̠ʐó/ 'to lay bricks'   /ɖʐó/ 'water'
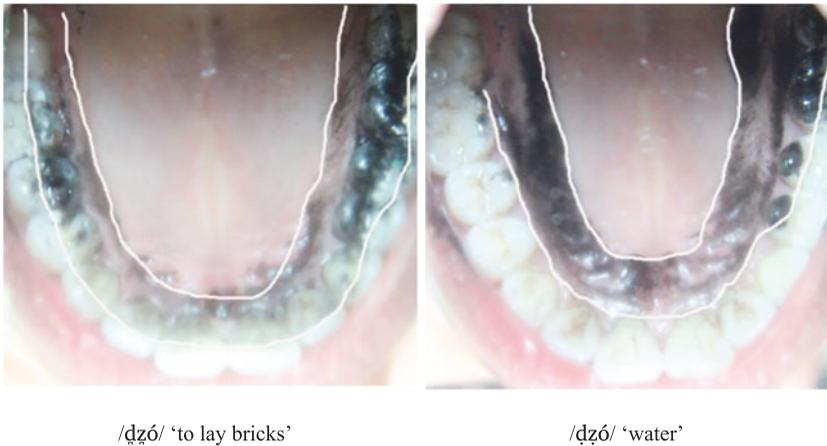
Figure 3 (Colour online) Palatograms of the words /d̠ʐó/ 'to lay bricks' and /ɖʐó/ 'water'.

and the roof of the mouth. Ersu retroflex affricates are typically single-contact trills. This is illustrated in Figure 6, with the words /ʈʂò/ 'gallbladder' and /ɖʐò/ 'pot, pan'. A contact, that is, a moment of closure of the oral cavity, is reflected on the spectrograms by a period of white space.

When retroflex affricates are followed by fricative vowels, the number of contacts between the tip of the tongue and the roof of the mouth may be increased to three. This is illustrated in Figure 7, with the words /ʈʂ$\acute{\gamma}$/ 'sweat' and /Nɖʐ$\acute{\gamma}$/ 'tile'.

Fricatives
Ersu contrasts fricatives at five places of articulation: (i) labiodental, (ii) dental, (iii) alveolar, (iv) alveolopalatal (palatalized laminal post-alveolar), and (v) velar. All but velar show a two-way contrast between voiceless and voiced: /f v/ (e.g. /fí/ 'mildew', /ví/ 'tinder'), /s̪ z̪/ (e.g. /s̪ɛ́/ 'air', /z̪è/ 'wife'), /s̠ z̠/ (e.g. /s̠ò/ 'blood', /z̠ó/ 'to scold; to be in debt'), /ɕ ʑ/ (e.g. /ɕó/ 'to sweep', /ʑó/ 'to soften (skin)'), /x/ (e.g. /xì/ 'bamboo').

The voiced labiodental fricative /v/ is contrastive with the voiced labial-velar approximant /w/ before /a/ and /o/, as in /vá/ 'net', /wá/ 'to be full, satisfied'. /f/ has a restricted distribution,

/d̪z̪ó/ 'water'   /dʐó/ 'to push'

Figure 4 (Colour online) Palatograms of the alveolar and alveolopalatal affricates in the words /d̪z̪ó/ 'water' and /dʐó/ 'to push'.

/ʈʐò/ 'gallbladder'   /ɖʐò/ 'pot, pan'
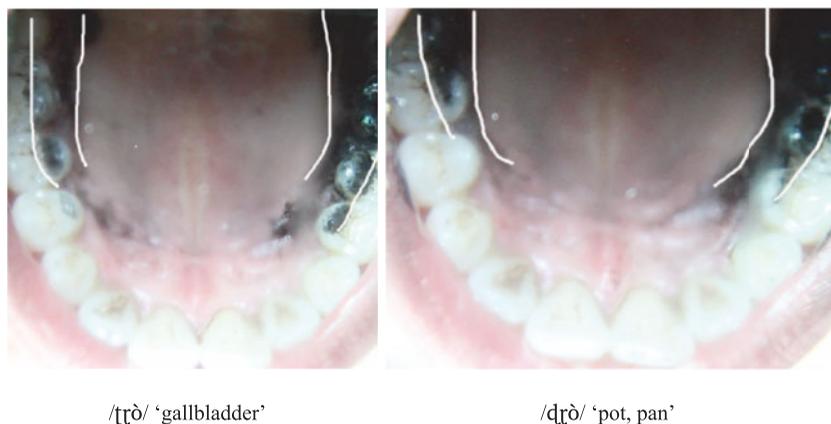
Figure 5 (Colour online) Palatograms of the words /ʈʐò/ 'gallbladder' and /ɖʐò/ 'pot, pan'.

occurring only before /i/ (as above) and /ɤ/ (as in /fɤ̃/ 'garlic'). In the latter environment, /f/ can be alternatively regarded as the allophone of /x/, as is also the case in SW Mandarin, with which Ersu is in close contact. Consider, for example, the Ersu word for 'kettle': /fɤ̃-fɤ́/, which is a loanword from SW Mandarin /fu²¹-fu⁴⁴/, corresponding to Standard Mandarin *hú* 壶 [xu³⁵].

Ersu exploits contrasts between grooved fricatives at the dental place of articulation and flat fricatives at the alveolar place of articulation. Figure 8 illustrates the three-way contrast between the dental, alveolar, and alveolopalatal places of articulation in fricatives with the words /z̪è/ 'wife', /ʐó/ 'to scold; to be in debt', and /ʑó/ 'to soften (skin)'.

Before /o/, /x/ has an allophone [ç], which we analyse phonemically as a sequence of /x/ and the palatal approximant /j/, e.g. /xjó/ [çó] 'to cry out' (see below on clusters with approximants).

Nasals
Ersu has nasals at three places of articulation: (i) bilabial (/m/), (ii) dental (denti-alveolar) (/n/) (see Figure 9), and (iii) velar (/ŋ/).

/ʈɽò/ 'gallbladder'   /ɖɽò/ 'pot, pan'
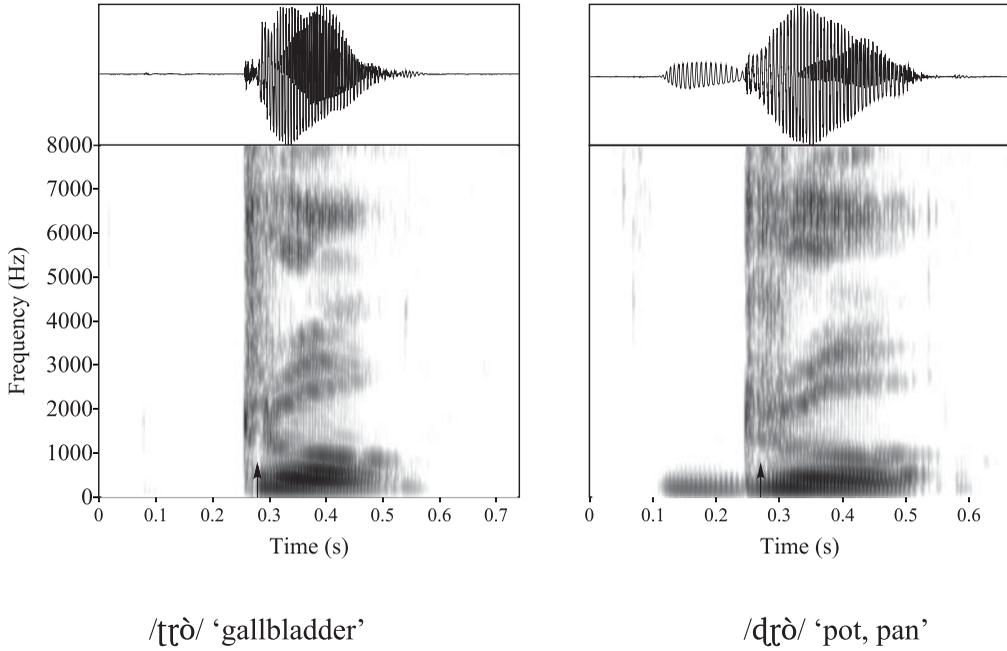
Figure 6 Waveforms and spectrograms of /ʈɽò/ 'gallbladder' and /ɖɽò/ 'pot, pan' (arrows indicate contacts between the tip of the tongue and the roof of the mouth).

/ʈɽv̩́/ 'sweat'   /Nɖɽv̩́/ 'tile'
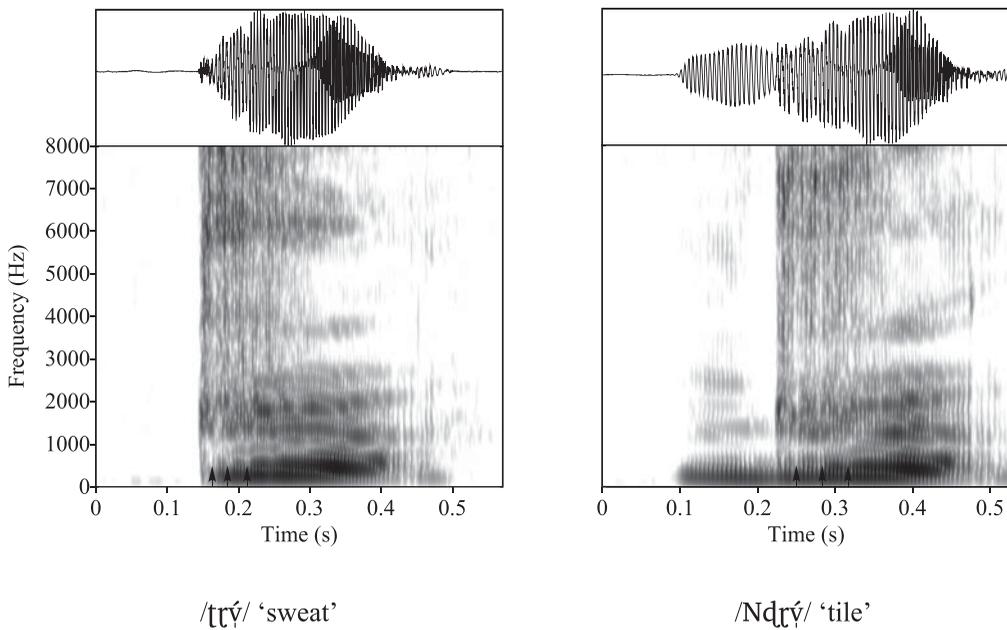
Figure 7 Waveforms and spectrograms of /ʈɽv̩́/ 'sweat' and /Nɖɽv̩́/ 'tile' (arrows indicate contacts between the tip of the tongue and the roof of the mouth).

/z̢ɛ̀/ 'wife'   /z̢ó/ 'to scold; to be in debt'   /z̢ó/ 'to soften (skin)'
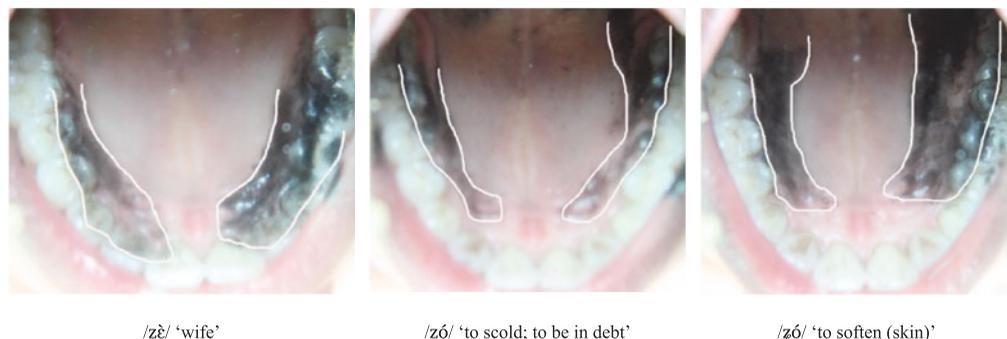
Figure 8 (Colour online) Palatograms of the words /z̢ɛ̀/ 'wife', /z̢ó/ 'to scold; to be in debt', and /z̢ó/ 'to soften (skin)'.
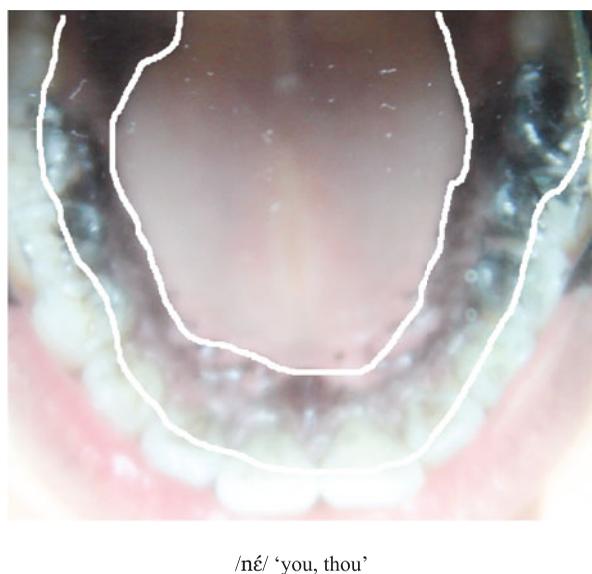
/nɛ́/ 'you, thou'

Figure 9 (Colour online) Palatogram of the word /nɛ́/ 'you, thou'.

Ersu has one syllabic nasal, /ŋ/, as in /s̢ź-ŋ́/ 'seven', /lwáŋ́ká/ 'ridge of a building', /ś̢ŋ́d̢z̢ɛ́/ 'satin, silk' (possibly from Chinese *língzi* 绫子 'damask silk', SW Mandarin /nin⁵³ts̢ɻ⁵³/).

Liquids

Ersu laterals are made with an occlusion in the alveolar region extending back to the back molars. This is illustrated in Figure 10, with the minimal pair /lá/ 'fertilizer, manure' and /ɬá/ 'month'.

Similar to retroflex affricates, the Ersu retroflex trill (/ʈ/) is produced with the point of contact on the roof of the mouth in the hard palate and with the tongue body braced against the sides of the teeth to allow for an aerodynamically induced movement of the tongue tip (see McGowan 1992: 2903, Spajić, Ladefoged & Bhaskararao 1996: 3; see Figure 11). Similar to the retroflex trill in Toda (Spajić et al. 1996: 13), the first contact of the tongue tip is made at the back of the alveolar ridge, whereas subsequent contacts are made slightly further forward near or at the alveolar ridge, so that the trill is realized as [ʈr].

/lá/ 'fertilizer, manure'   /ɬá/ 'month'
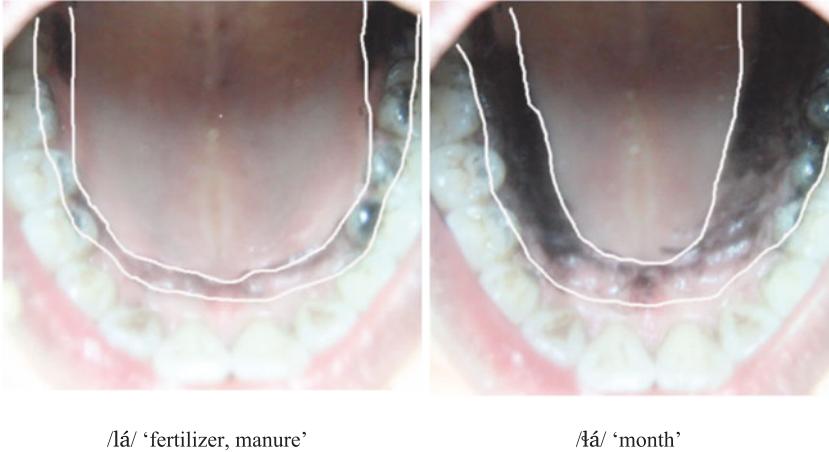
Figure 10 (Colour online) Palatograms of the words /lá/ 'fertilizer, manure' and /ɬá/ 'month'.
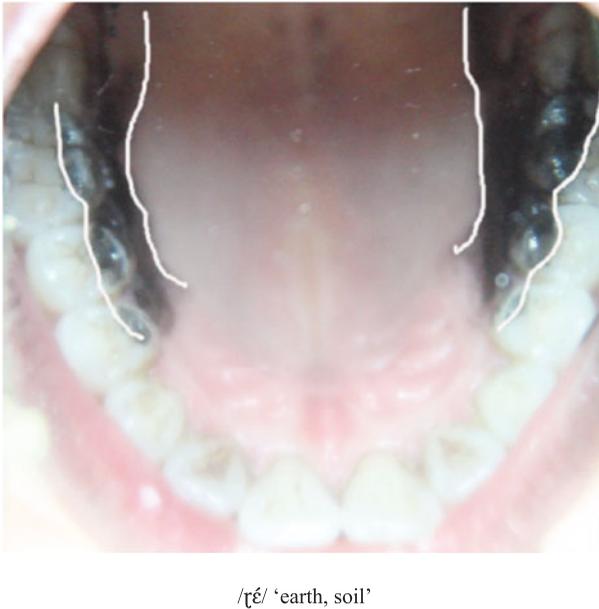
/ʐé/ 'earth, soil'

Figure 11 (Colour online) Palatogram of the word /ʐé/ 'earth, soil'.

Ersu trill typically has two contacts. However, similar to trilled retroflex affricates, a trill followed by a fricative vowel may have a larger number of contacts (four to five). This is illustrated in Figure 12.

Clusters
Ersu has a rich inventory of clusters, including (i) clusters with approximants, (ii) prenasalized clusters, and (iii) clusters with a schwa-like segment.

Clusters with approximants
The approximants /w/ and /j/ may occur after a broad range of initials and may be realized as secondary labialization or palatalization of the first position consonant. Of the two

/ʈá/ 'chicken'   /ʈó/ 'bone; horn'   /ʈʐ́/ 'to laugh'
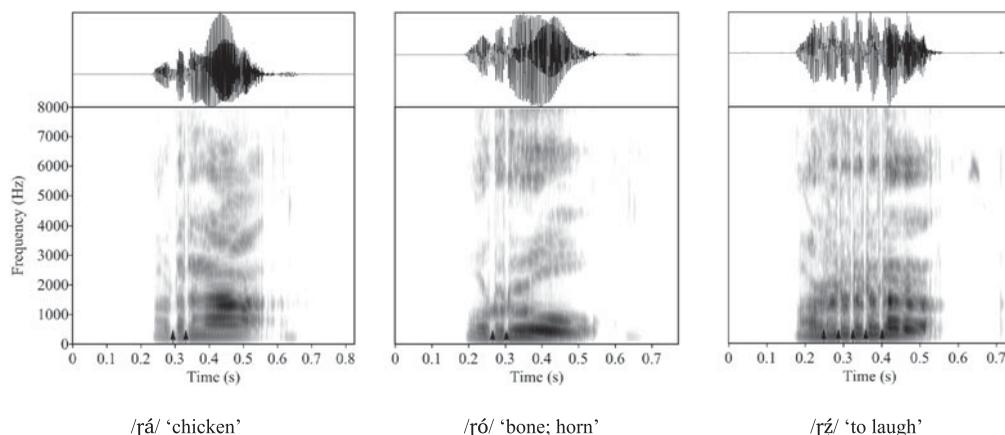
Figure 12 Waveforms and spectrograms of the words /ʈá/ 'chicken', /ʈó/ 'bone; horn', and /ʈʐ́/ 'to laugh' (arrows indicate contacts between the tip of the tongue and the roof of the mouth).

approximants, /w/ has the broadest distribution, occurring after bilabial, dental, and velar stops, dental affricates, and laterals. However, most of these clusters with /w/ can only be followed by /a/ (as in /kwá/ 'to take off', compare to /ká/ 'to be stupid'). In addition, clusters with velar stops and /w/ can also be followed by /a˞/ (as in /nà-kwá˞/ 'to put inside', compare to /d̪è-ź̪-kà˞/ 'to bear a grudge').

/j/ has a more restricted distribution, occurring only after (bilabial and dental) nasals and laterals, and it can only be followed by the vowels /o/ and /a/. (/j/ may also occur after /x/, which combination can only be followed by /o/. The cluster /xjo/ is realized as [ç], as in /xjó/ [çó] 'to cry out', see above). Consider the following (near) minimal pairs: /má/ 'mother, female' vs. /mjábó/ 'tear'; /nà-ná/ 'to occupy' vs. /njá/ 'child'; /lá/ 'fertilizer, manure' vs. /ljá/ 'to paint'; and /Nts̪ʰòɬó/ 'flea' vs. /mé-ɬjó/ 'lightning'.

It is important to note that the realization of /w/ and /j/ in clusters ranges between a separate segment with a clear segmental boundary (mostly before /a/) and a segment with a secondary articulation, that is, a segment with a lesser degree of stricture that accompanies a primary articulation of a higher degree (mostly before /o/) (see Ladefoged & Maddieson [1996](#): 354). Examples include /njá/ [njá] 'child', /ljá/ [ljá] 'to paint', /njó/ or /nʲó/ [ɲó] 'day', /ljò/ or /lʲò/ [ʎò] 'arm spread (measure of length)'. We note that the addition of the lip rounding gesture (in the case of /w/) and raising of the body of the tongue (in the case of /j/) have a strong acoustic effect on both the preceding consonant and the following vowel. Overall, compared to clusters with /w/ and /j/ in the closely related Lizu and Duoxu languages, Ersu medials /w/ and /j/ exhibit stronger assimilatory influence on neighboring segments (see Chirkova & Handel [2013](#)b for a detailed discussion).

In light of the above, Ersu alveolopalatal affricates and fricatives, which only occur in native vocabulary before the vowels /i o a/, can be alternatively analysed as palatalized allophones of dentals, alveolars and/or velars, followed by the high front vowel or the palatal approximant /j/.

Prenasalized clusters

Prenasalization in Ersu is contrastive and occurs not only before voiced stops and affricates, but also before voiceless aspirated ones. Prenasalized stops and affricates are found in all places of articulation. The place of articulation is always homorganic with that of the obstruent in the cluster. Thus, we use the archiphoneme N to refer to the homorganic nasal in prenasalized clusters. Compare the contrast between plain onsets in the following minimal pairs: /bò/ 'string' vs. /Nbò/ 'horse', /pʰó/ 'to escape' vs. /Npʰó/ 'to steal', /d̪z̪à/ 'fodder' vs. /Nd̪z̪à/

/pʰó/ 'to escape'   /Npʰó/ 'to steal'

/tʰwá/ 'mule'   /Ntʰwá/ 'to be sharp'
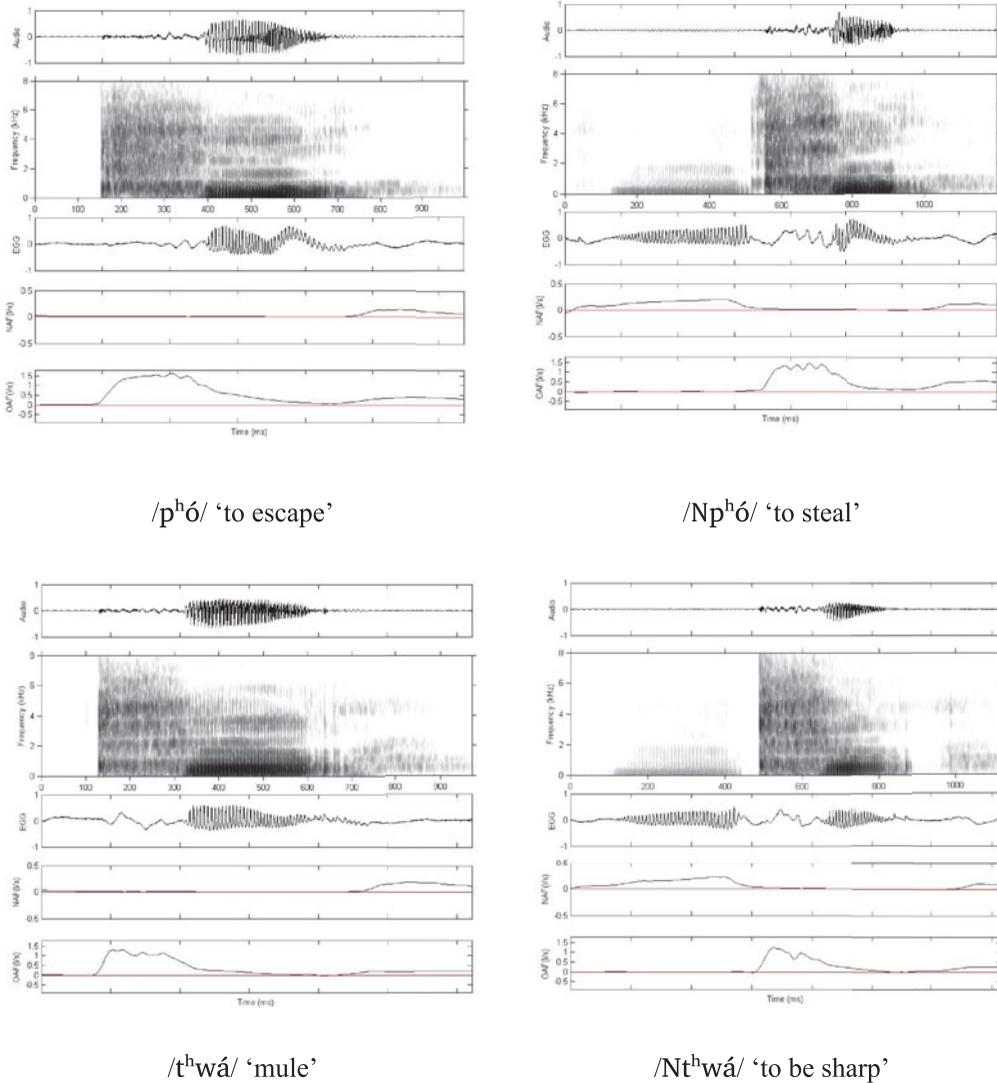
Figure 13 (Colour online) Audio waveforms, spectrograms, electroglottograph waveforms (EGG), nasal airflows (NAF), and oral airflows (OAF) for the pairs /pʰó/ 'to escape' vs. /Npʰó/ 'to steal' and /tʰwá/ 'mule' vs. /Ntʰwá/ 'to be sharp'.

'Chinese people', /ʈ͡ʂʰá/ 'hot' vs. /Nʈ͡ʂʰá/ 'mark, sign', /ɖɻò/ 'pot, pan' vs. /Nɖɻò/ 'dirt, filth', /(mélí) té ʈɻʰó/ 'one piece (of land)' vs. /té Nʈɻʰó/ 'one handful (e.g. of rice)'.

In prenasalized voiceless aspirated stops and affricates, we observe regular glottal pulsing during the nasal section but not during the voiceless stop, as detailed in Figure 13. This figure shows that the nasal section is produced with a complete closure within the oral tract (no oral airflow) and with air moving through the nasal cavity (nasal airflow for the entire duration of the segment). It is therefore a characteristic nasal.

Clusters with a schwa-like segment
Ersu has seven voiceless (unaspirated) stops, affricates, and fricatives (/əp ət ək əʈ͡ʂ əʦ əʨ əx/) and seven voiced stops, fricatives, and nasals (/əb əd əg əʐ əm ən əŋ/) that can be preceded

/ət̠s̠ɛ́/ 'interest'   /t̠s̠ɛ́/ 'cloud'
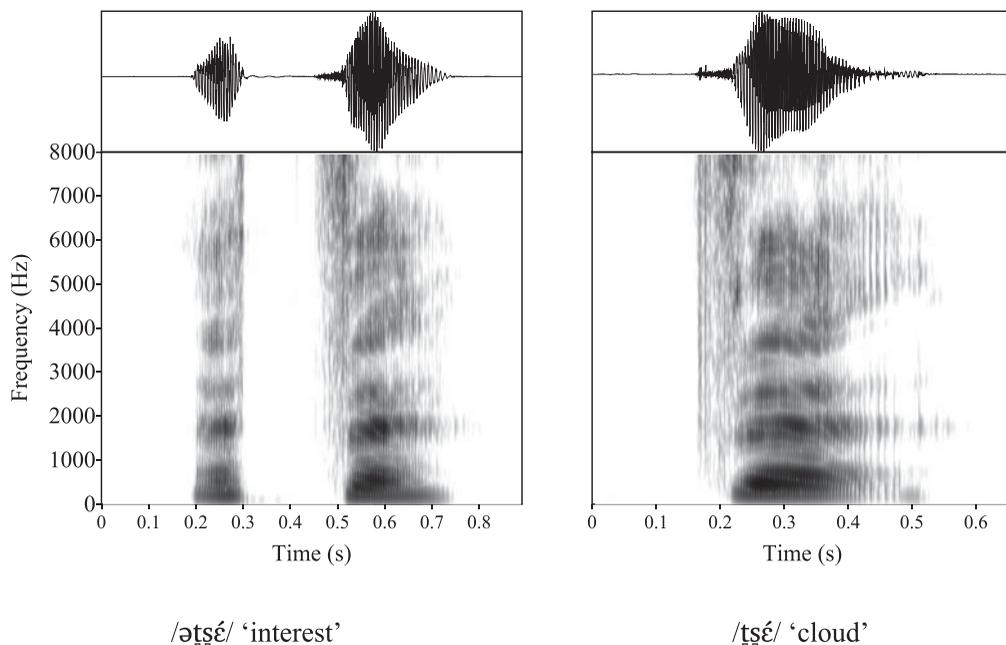
Figure 14 Waveforms and spectrograms for the minimal pair /ət̠s̠ɛ́/ 'interest' vs. /t̠s̠ɛ́/ 'cloud'.

by a segment with a schwa-like formant structure. Compare the minimal pairs /ət̠s̠ɛ́/ 'interest' vs. /t̠s̠ɛ́/ 'cloud', and /əné/ 'to be heavy; to be deep' vs. /né/ 'two' in Figures 14 and 15.

Clusters with a schwa-like segment can be tentatively identified with earlier consonant clusters. Independent evidence for such clusters can be found in Ersu words that etymologically correspond to Proto-Lolo-Burmese and Proto-Tibeto-Burman forms with initial clusters (see Bradley 1979: 144, 1985: 242; Matisoff 2003:37). For example, Ersu /əné/ 'to be deep' corresponds to Proto-Loloish *ʔ-nak[L] 'deep' (Matisoff 2003: 37) (see Chirkova & Handel 2013a for a detailed discussion). Clusters with a schwa-like segment are in the process of disappearing from this language, merging with corresponding simple onsets.[4]

In addition to clusters with a schwa-like segment, Ersu also marginally has preaspirated clusters, which are restricted to loanwords from Tibetan. Examples include /htóNbá/ 'to be empty' (WT *stong pa*), /hkwàɽá/ 'to turn, to circle' (WT *skor ba*). In these words, preaspiration diachronically derives from stop clusters with the preradical *s-* in Old Tibetan. The acoustic quality of Ersu preaspirated clusters is different from voiceless segments preceded by a schwa-like segment. The main differences include (i) the longer duration, and (ii) the clear formant structure of the schwa-like segment, as compared to preaspiration. These differences are illustrated in Figure 16, with the pair /ətó/ 'to jump' vs. /htóNbá/ 'to be empty'.

---

[4] We note that Ersu nasals preceded by a schwa-like segment correspond to what we impressionistically call geminated nasals in the closely related Lizu language. Our principal Lizu language consultant marginally distinguishes between, according to his own analysis, a relatively shorter nasal initial and a relatively longer ('geminated') nasal initial in a few minimal pairs, such as /[F]ne/ 'you, thou' vs. /[F]ne/ [[F]nne] 'deep'. A more systematic comparison between Ersu words with this type of clusters and their cognates in Lizu will be undertaken in forthcoming fieldwork.

/ənɛ́/ 'to be heavy; to be deep'   /nɛ́/ 'two'
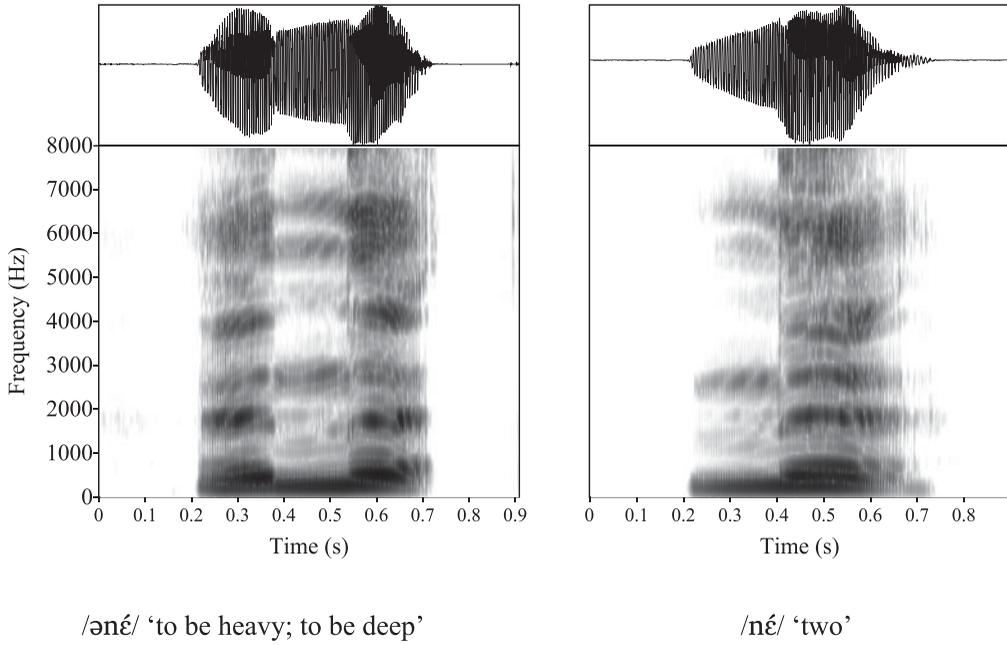
Figure 15 Waveforms and spectrograms for the minimal pair /ənɛ́/ 'to be heavy; to be deep' vs. /nɛ́/ 'two'.

/ətó/ 'to jump'   /htóNbá/ 'to be empty'
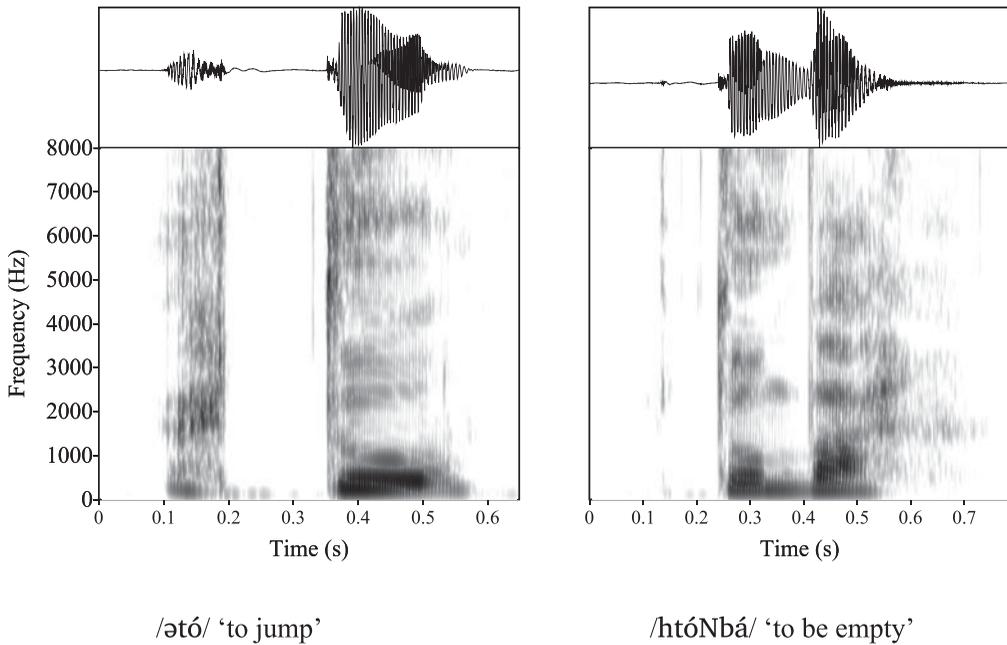
Figure 16 Waveforms and spectrograms for the pair /ətó/ 'to jump' vs. /htóNbá/ 'to be empty'.

Vowels

Ersu has eight vowel phonemes, of which four are plain (/i ɛ a o/), two are fricative (/z̩ ʋ̩/), and two are rhotacized (/ɚ a˞/). See the vowel chart plotted on the relative F1/F2 formant values.
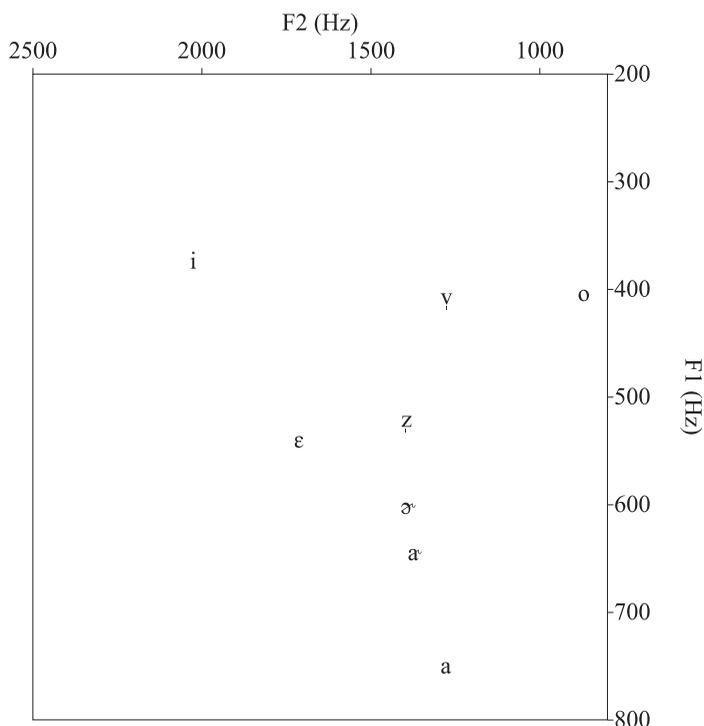
i	íz̩á	'son'	pí	'dregs'	ts̩í	'hair'
ɛ	élá	'ram'	pɛ́	'to give as a present'	ts̩ɛ́	'cloud'
a	á	'I'	pá	'place of provenance'	ts̩á	'to tie'
o	ódz̩á	'pear'	pó	'wrapping'	ts̩ó	'grain cabinet'
z̩	z̩̀	'snow'	pź̩-pź̩	'to be flat'	ts̩ź̩	'to feed'
ʋ̩	ʋ̩́	'wine'	pʋ̩́	'potato'	ts̩ʋ̩́	'hearth'
a˞			pà˞	'grain'		
ɚ	ɚ́	'year'	pɚ̀ [pz̩̀]	'thin rope'		

Plain vowels

/i/ has a fairly broad range of realizations. It ranges from [i] (essentially in loan vocabulary from SW Mandarin and Nuosu) to [je̞] (mostly after bilabial initials) and [e] (mostly after dental, alveolar, and alveolopalatal initials). The value range for F1 is between 264 and 347 Hz, and for F2 between 1920 and 2274 Hz. Examples include: /ɕíkwá/ [ɕíkwá] 'watermelon' (Chinese *xīguā* 西瓜, SW Mandarin /ɕi⁴⁴kua⁴⁴/), /ítṣʰʋ̩́/ [éts̩ʰʋ̩́] 'ladle, a long handled spoon, generally made of wood' (Nuosu *it chyp* [i⁵⁵tʂʰɻ̩²¹]);[5] /pí/ [pjé̞] 'dregs', /mí/ [mjé̞] 'monkey', and /s̩í/ [s̩é] 'wood' (compare to /s̩é/ 'air').

---

[5] Nuosu examples are from the online Nuosu Yi-Chinese-English Glossary (http://www.yihanyingcihui. net/?page_id=6065&lang=en), accessed January 2013. In the transcription system of Nuosu, syllable-final consonant symbols (e.g. *t* and *p* in the word for 'ladle') are used to mark tone (respectively, the high level, 55, and the low falling tone, 21).

After alveolar affricates, dental and alveolar fricatives, and velar stops, /ɛ/ may be realized as retroflex (characterized by lowered F3 values). For example, in the sound files compare the realization of the word /ʑé/ 'to crawl, to climb' in isolation and the three repetitions of that word in the compound /Nbí ʑé/ 'to climb mountains'.

Fricative vowels

Fricative vowels (Ersu /z̩/ and /v̩/) are defined as vowels that are produced with the tongue in essentially the same position as in the corresponding fricatives (Ladefoged & Maddieson 1990: 117, 1996: 314). The constriction of the tongue tip or lips produces alveolar and labiodental frication, respectively.

The two fricative vowels in Ersu (/z̩ v̩/) are independent phonemes that co-occur with a broad range of initials. Therefore, they are distinct from the known cases of fricative vowels in Mandarin or Nuosu (see e.g. Chao 1972 [1948]; Li & Ma 1983: 36; Ladefoged & Maddieson 1996: 314), where syllabic fricatives can be viewed as conditioned variants of other (high) vowels.

Ersu fricative vowels display periodic vocal fold vibration and clear formant structure, as is typical of vowels. The two are differentiated by the configuration of the lips: spread for /z̩/ and rounded for /v̩/. /v̩/ is produced with a pronounced lip compression, whereby the lower lip is raised, while the upper lip remains in a static position (as characteristic for the articulation of labiodental fricatives, see Laver 1994: 250). Examples include /z̩̀/ 'snow', /v̩́/ 'wine', /z̩ź̩/ 'shoe', and /z̩v̩́/ 'oil' (see also video clips 'shoe' and 'oil'). Ersu fricative vowels are accompanied by fricative noise: the high-frequency energy noise in the 3000–6000 Hz region for /z̩/, and a relatively flat spectrum for /v̩/ (as typical of labiodental fricatives) (see Ladefoged & Maddieson 1996: 173–176). The fricative noise is more diffuse and weaker in intensity than that found for other fricatives (see Figure 17).

The vowel /z̩/ may occur after bilabial stops, dental and alveolar affricates and fricatives, retroflex affricates, /ɻ/, and, in a very few cases, also after the voiced velar initials /g Ng ŋ/. It has a broad range of realization:

(i) After bilabial stop initials, /z̩/ is realized as [zz̩] after non-aspirated initials and as devoiced, [s̩], after aspirated stop initials. Examples include /bz̩̀/ [bzz̩̀] 'bee', /pź̩-pź̩/ [pź̩-pzź̩] 'to be flat', /pʰz̩̀pó/ [ps̩̀pó] 'wood shavings'.

(ii) After dental and alveolar affricates and fricatives, retroflex affricates and /ɻ/, /z̩/ is realized as homorganic to the preceding consonant onset. Examples include /sz̩̀/ 'matter, affair' (Chinese *shì* 事, SW Mandarin /sⁿ²¹³/), /zz̩̀/ 'shoe', /tsź̩/ 'to feed', /ts̩ʰz̩̀/ 'salt', /sz̩̀/ [sṣ̩̀] 'meat', /zź̩/ [zz̩] 'to wear, to put on', /ʈʂź̩/ [ʈʂɻ̩́] 'star', /ʈʂʰź̩/ [ʈʂʰɻ̩́] 'to cut with a sickle', /ɖʐź̩/ [ɖʐɻ̩́] 'glutinous rice', /ɻź̩/ [ɻɻ̩́] 'to laugh'.

(iii) After velar initials, /z̩/ is realized close to [ɤ], as in /gź̩ɻóətsź̩/ [gɤ́ɻóətsź̩] 'spine, backbone' (compare to /gv̩́/ 'boat', /gó-gó/ 'light (ADJ)').

/v̩/ has a broader distribution and may occur after bilabial and velar stops; dental, alveolar and retroflex affricates; nasals; and dental and alveolar fricatives. It may be realized as [v̩] or [vʋ] in free variation (compare the two realizations of the word /gv̩́/ 'boat'). Similar to /z̩/, /v̩/ has a fairly broad range of realization:

(i) After bilabial stop initials and retroflex affricates, /v̩/ is realized close to [ʙ]. This is similar to its realization in the closely related Lizu language. Examples include /pv̩́/ [pʙ̥] 'potato', /tʰɛ̀-pʰv̩́/ [tʰɛ̀-pʰʙ̥] 'to change, to turn into, to transform', /bv̩́/ [ʙ̥] 'wild cat', /ʈʂv̩́/ 'sweat', /ʈʂʰv̩́/ 'six', /Nɖʐv̩́/ 'tile' (see also Figure 7 above and video clips 'potato', 'to change, to turn into, to transform', 'wild cat'). The bilabial trill is particularly evident in the minimal contrastive pair /bv̩́/ [ʙ̥] 'wild cat' vs. /bó/ [bó] 'to have, to possess'.

Figure 17 Acoustic spectra and waveforms and spectrograms of the words /ʐ̩/ 'snow', /ɣ̩́/ 'wine' (an arrow indicates high-frequency noise in the 3000–6000 Hz region).

(ii) After /m/, /ɣ/ is realized as a voiced syllabic bilabial nasal ([m̩]). For example, /mɣ̀tʂɛ́/ [m̩tʂɛ́] 'carpenter', /mɣ́tʂʐ̩/ [m̩tʂʐ̩] 'cat'.[6]

Rhotacized vowels
Ersu has two rhotacized vowels: /ɚ/ (as in /ɚ́/ 'ashes; year; to be white; to bark', /xɚ̀mìNtɕʰí/ 'south', reportedly from Nuosu (*yyx*) *hmy* [(ʐ̩³⁴)m̩³³] 'south') and /aʵ/ (as in /vaʵ́/ 'slave',

---

[6] We note that this distribution has a close parallel in most Nuosu varieties, where the voiced syllabic bilabial nasal [m̩] is the usual allophone of a syllable with a voiced or voiceless bilabial nasal initial followed by /u/ (Bradley 2008). Furthermore, similar to Ersu, syllables with the high vowels /i/ and /u/ are an environment conducive to the development of syllables with unusual phonetic forms in many Ngwi languages (such as Adur, various Nuosu or Liangshan Yi varieties, and Sanie, see Pan 2001, Matisoff 2006, Bradley 2008) as well as some Qiangic languages, such as Namuzi (Lama Ziwo 1994). These unusual forms include syllabic fricatives or trills after initial bilabial and velar stops, and labial-velar nasals after initial bilabial and velar nasals.

Figure 18 Spectrograms of /vá/ 'net' vs. /vǽ/ 'slave' (arrows indicate F3).

/xǽ/ 'bear; needle'). Compared to their oral counterparts (/z̩/ and /a/, respectively), /ɚ/ and /æ˞/ have a lowered frequency of the third formant. This is illustrated in Figure 18, with the minimal pair /vá/ 'net' vs. /vǽ/ 'slave'.

/æ˞/ has a broad distribution and co-occurs with bilabial and velar stops, /m/, /ŋ/, /v/, and /x/. (After bilabial initials, /æ˞/ may be realized as the sequence [ɹa], e.g. /bǽ/ [bɹá] 'to be full', /mǽ/ [mɹá] 'to sleep'.) Conversely, /ɚ/ mostly occurs in isolation (as in /ɚ́/ 'ashes; year; to be white; to bark'). In addition, /ɚ/ also occurs after bilabial stops and /x/ (as in /xɚ̂mÌNtɕʰɪ́/ 'south'). After bilabial stops /ɚ/ is realized with frication. This is similar to the realization of /z̩/ after bilabial stops. However, in contrast to /z̩/, /ɚ/ is pronounced with the tongue curved in a convex shape. Examples include /bɚ́/ [bź̩] 'crown of a head', /pɚ̀/ [pz̩̀] 'thin rope', /pʰɚ́/ [pʂ̩́] 'Tibetan'.

Overall, Ersu can be said to have an unbalanced vowel system in that it has a high front vowel /i/ without a corresponding high back vowel /u/; a back close-mid vowel /o/ without a corresponding front close-mid vowel /e/; and an open-mid low vowel /ɛ/ without a corresponding back open-mid vowel /ɔ/. In our analysis, this may be due to an ongoing realignment of the vowel system of Ersu, following the development of the fricative vowels /z̩/ and /v̩/ from the high vowels /i/ and /u/ in this language. The development of the fricative vowels from /i/ and /u/ is suggested, on the one hand, by the synchronic distributional evidence, considered in the light of aerodynamic constraints outlined in Ohala (1983) and, on the other hand, by comparative evidence from the closely related Lizu and Duoxu languages. We note that Ersu /z̩/ does not co-occur with (dental and velar) stop initials, but it co-occurs with affricate initials instead. The vowel /v̩/ does not co-occur with dental stop initials, but it co-occurs with alveolar affricate initials. This complementary distribution can be explained as an outcome of sound change whereby dental and velar stops developed an affricated release when followed by high vowels. This is due to the fact that the high velocity of the airflow created upon release of a stop lasts longer when the stop precedes a close vowel as opposed to an open vowel (Ohala 1983: 204–205). From a comparative perspective, Ersu /z̩/ has multiple correspondences in Lizu and Duoxu, including /i/ (as in 'bee': Ersu /bź̩/, Lizu

/^{R}bi/, Duoxu /bi^{31}/; 'shoe': Ersu /z̩ź/, Duoxu /ʐi^{33}/), /e/ (as in 'hair': Ersu /tʂí/, Lizu /^{F}tɕe/) as well as non-high vowels preceded by the palatal approximant /j/ (as in 'mountain': Ersu /Nbí/ [Nbjé], Lizu /^{R}Nbje/) (see Chirkova & Handel 2013b). These correspondence patterns reveal complex developments, which contribute to a realignment of the vowel system of Ersu, whereby the earlier phoneme /e/ is moving into the vacated /i/ space.

Nasalized vowels

Ersu marginally has a set of nasalized vowels. Nasalized vowels are generally restricted to recent loanwords from Mandarin Chinese, where the donor language has the nasal codas /n/ or /ŋ/ (as in /kã̀/ 'steel', Chinese *gāng* 钢, SW Mandarin /kaŋ^{44}/). (In addition, nasalized vowels are attested in two native Ersu words in our corpus, /õ̃ã́/ 'goose' and /è̃ã́/ 'duck'.) For that reason, vowel nasalization in Ersu must be regarded as subphonemic, and only needs to be marked in those cases where it is unpredictable (i.e. in recent loanwords). It is interesting to note that in older loanwords, where the original nasal coda is followed by a syllable that begins with a vowel or a nasal or when it is word-final, the original nasal element is in most cases lost without compensation, as in /tʂwá/ 'brick' from *zhuān* (SW Mandarin 砖 /tʂuan^{44}/), /pètì/ 'silver' from *báidìng* 白锭 'white ingot' (SW Mandarin /pei^{21}tin^{213}/). This is similar to the situation in the closely related Lizu and Duoxu languages (Chirkova & Chen 2013, Chirkova & Handel 2013b).

Vowel harmony

In disyllabic domains, we observe regressive vowel assimilation. The vowel qualities can be divided into two sets: (i) the low vowels /a aɻ/, and (ii) the remaining, non-low vowels, that is, /i ɛ o z̩ v̩/. Vowel harmony appears to only apply to directional prefixes and the number 'one' (that is, it is restricted to high frequency morphemes). Consider expressions consisting of the numeral 'one' (/té/ in isolation) followed by various nouns: /tá z̩á/ 'one hundred', /tá ká/ 'one strip', /Ntʂʰè̀ tà pá/ 'one grain of rice', /té pʰó/ 'one set (of clothing)', /té pv̩́/ 'one tree; ten cents', /té dz̩ź/ 'one sentence'. Like many languages that display vowel harmony in polysyllabic lexical items Ersu has two forms for affixes, such as verbal directional affixes, e.g. /kʰá-lá/ 'to come in (in the direction to the speaker)' vs. /kʰɛ́-jí/ 'to enter (in the direction away from the speaker)'.

Syllable structure

The canonical Ersu syllable minimally consists of an obligatory nucleus and a tone. It may also contain up to three optional elements in the following linear structure: (C1)(C2)(C3)V, where C1 can be nasal (/N/) or a schwa-like segment (/ə/); C2 can be any consonant; C3 can be either /w/ or /j/; and V stands for vowel, and parentheses indicate optional constituents. Zero-initial words can be preceded by a non-phonemic glottal stop (e.g. /õ̃ã́/ [ʔõ̃ã́] 'goose', /ódz̩á/ [ʔódz̩á] 'pear').

(1)   V   /á/   'I, first person singular pronoun'
(2)   CV   /tá/   'flag' (WT *dar*), /tʰá/ 'millstone', /bò/ 'string'
(3)   CCV   /ətá/   'wether, castrated ram', /tʰwá/ 'mule', /Nbò/ 'horse'
(4)   CCCV   /Ntʰwá/   'to be sharp'

In addition, (recent) loanwords may have the following structures:

(5)   CVŋ   /píẗʰàŋ/   'crystal sugar'
                             (from Chinese *bīngtáng* 冰糖, SW Mandarin /pin^{44}tʰaŋ^{21}/)

(6)   CVV(V)   /fɛ́ẗʰiào/   'rice noodle'
                             (from Chinese *fěntiáo* 粉条, SW Mandarin /fen^{53}tʰiao^{21}/)

Loanwords from Tibetan may also have /h/ in the C1 slot, as in /ɦkwàɻá/ 'to turn, to circle'.

Similar to its linguistic neighbours, Ersu is phonologically monosyllabic with a strong tendency towards disyllabicity in its lexicon. Trisyllabic and quadrisyllabic words are mostly composite, e.g. /lɛ́máká/ 'thumb' (< /lɛ́/ 'hand'), /ʂɤ́Nbɤ́ əpɛ́əkɤ́/ 'nostril' (< /ʂɤ́Nbɤ́/ 'nose', /əpɛ́əkɤ́/ 'hole, cavity'), although a handful of trisyllabic monomorphemic words (both native and loanwords) do exist (e.g. /xə̀mìNtɕʰí/ 'south').

In disyllabic composite forms, where the second syllable has zero initial, the two adjacent vowels merge into one vowel or a diphthong, a process that typically results in a tone change. This change characteristically occurs when the perfective marker /á/ (which has an etymological high-register tone) is added to a verb stem. For example, compare the realization of the verb /dè-pʰwá/ [dà-pʰʷá] 'to smash' in isolation and when followed by the perfective marker /á/, i.e. /dè-pʰwà-á/ 'have smashed'.

The diminutive morpheme /ji/ often fuses with the preceding vowel resulting in a diphthong that combines the original vowel with the offglide [j]. Examples include 'armpit', which is /jíbá-jì/ in careful pronunciation and /jíbáj/ in a more rapid speech tempo (note a clear falling tone contour on the second syllable of the fused form, resulting from the low tone of /jì/ being conjoined with the high tone of /bá/); and 'lamb', which is /lábʑ́-jì/ in careful pronunciation and /lábʑ́j/ in a more rapid speech tempo.

Prosodic organization

Ersu is a register tone language with two registers: Low and High (hereafter L and H), thus bearing resemblance to the tonal system of Tibetan (see e.g. J. Sun 1997). In polysyllabic domains, there is tone reduction in non-initial syllables, resulting in highly restricted tone patterns in polysyllabic words.

Monosyllabic words and compounds

The register contrast in Ersu can be exemplified with the following minimal pairs: /bɤ̀/ 'to plough' vs. /bɤ́/ 'wild cat', /là/ 'musk deer' vs. /lá/ 'fertilizer, manure', /Ndʐò/ 'dirt, filth' vs. /Ndʐó/ 'leather, skin', /tʂʰò/ 'dog' vs. /tʂʰó/ 'sound, melody'. While the register contrast is fundamental in this language, surface pitch contours are subject to variation in both registers (as illustrated in Figure 19).

We note that monosyllabic words that are rarely used in isolation (such as verbs and measure words) and loanwords are often realized in the context of elicitation in the high register, which we, for that reason, regard as the unmarked register in this language.

Polysyllabic words and compounds

In disyllabic domains (both compounds and composite lexical words), the domain-initial syllable retains its tone whereas the non-initial syllable does not. Disyllabic words and compounds that begin with a morpheme or a word in the high register are invariably realized with H tone on both syllables. Conversely, in the case of disyllabic words and compounds that begin with a morpheme or a word in the low register, two tonal patterns are possible: (i) L tone on both syllables, and (ii) L tone on the domain-initial syllable and H on the second syllable. The two patterns are in free variation, as is the case in the word /ɣ̀ɽà/ 'cloth': [vʋ^{33}ɽra^{44}] or [vʋ^{33}ɽra^{31}]. Examples of the tonal patterns on disyllabic domains include (i) the compounds /ŋwà-dʑí/ 'cow shed' and /Nbò-dʑí/ 'horse pen', which begin each with a L register word (/ŋwà/ 'cow' and /Nbò/ 'horse', respectively; the word 'shed, pen' has an etymological low-register tone, i.e. /dʑì/); and (ii) the compounds /vɛ́-dʑí/ 'pig shed' and /jó-dʑí/ 'sheep pen', which begin each with an H register word (/vɛ́/ 'pig' and /jó/ 'sheep', respectively).

The same two patterns are attested in composite lexical words. Examples of minimal pairs for the two tonal patterns include: /Ndʑó-Ndʑó/ 'to make friends', /Ndʑò-Ndʑó/ 'ear of millet' (see Figure 20); /njó-njó/ 'milk; breast', /njò-njó/ 'to be soft'; /tɕó-tɕó/ 'to wrap, to bind', /tɕò-tɕó/ 'maternal uncle' (Chinese *jiùjiu* 舅舅, SW Mandarin /tɕiəu^{213}tɕiəu^{44}/); /kʰɛ́-jí/ 'to enter', /kʰɛ̀-jí/ 'to go live (inside)'.
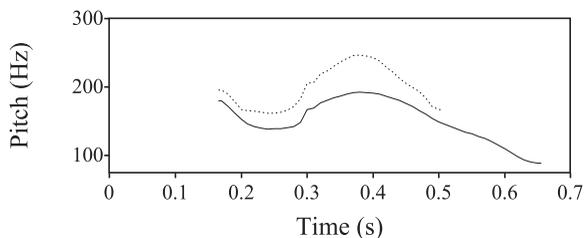
/bɣ̀/ 'to plough' and /bɣ́/ 'wild cat'
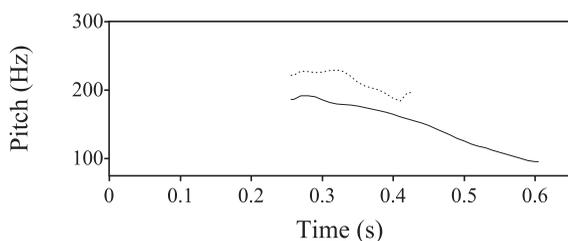
/ʈʂʰò/ 'dog' and /ʈʂʰó/ 'sound, melody'

Figure 19 Pitch contours of the low and high register tones by a male speaker illustrated with the minimal pairs /bɣ̀/ 'to plough' vs. /bɣ́/ 'wild cat', and /ʈʂʰò/ 'dog' vs. /ʈʂʰó/ 'sound, melody'. The low tone contour is represented by a solid line, the high tone contour is represented by a dotted line.

/Ndʐó-Ndʐó/ 'to make friends'
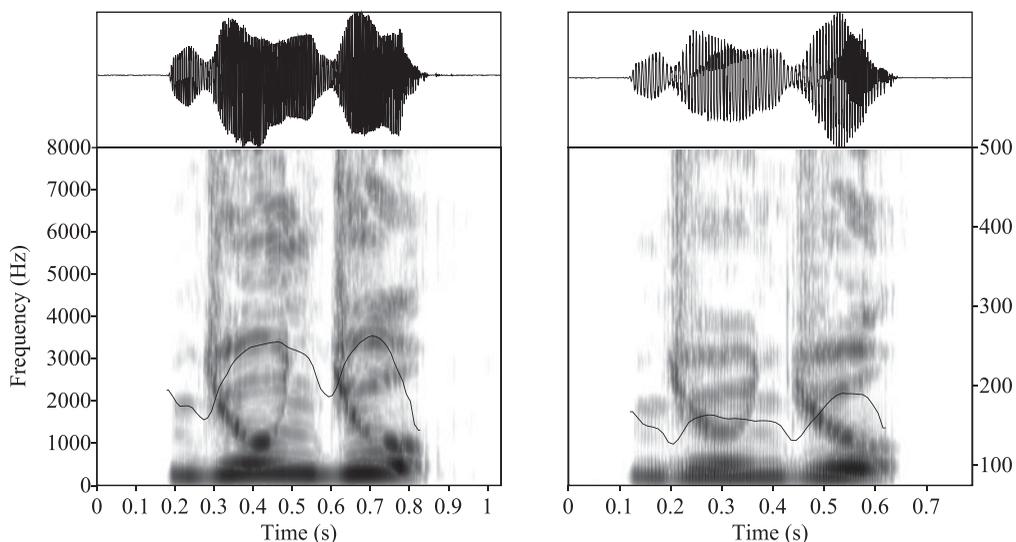
/Ndʐò-Ndʐó/ 'ear of millet'

Figure 20 Illustration of the H vs. L contrast in disyllabic items /Ndʐó-Ndʐó/ 'to make friends' and /Ndʐò-Ndʐó/ 'ear of millet' by a male speaker.

In addition, Ersu disyllabic words marginally have a third tonal pattern, in which the high f0 peak is realized before the end of the first syllable, where the pitch starts to fall already and continues to fall in the second syllable. Perceptually, the second syllable of these words sounds much less prominent than the first syllable, giving rise to the impression that it is unstressed. This tonal pattern is essentially observed in Chinese loanwords, in fusions as well as in a handful of native Ersu words. Examples include /fɛ́tʰiào/ 'rice noodle', /pít�ǹ/ 'crystal sugar', /tʂʰálà/ 'tomb', /ṣɛ́tɛ̀/ 'who'.

The same three tonal patterns are observed in trisyllabic compounds. These include:

(i)   H tone on all syllables within the domain, for those domains that begin with an H tone word (HHH), as in:

/zɤ́/ 'fish' + /mɛ́Ntʂʰɛ́/ 'tail' > /zɤ́ mɛ́Ntʂʰɛ́/ 'fish tail'
/jó/ 'sheep' + /Ndʐɤ̀pí/ 'skin, leather' > /jó Ndʐɤ́pí/ 'sheep skin'

(ii) L tone on all syllables of the domain, for those compounds that begin with an L tone word (optionally, with the H tone on the last syllable of the domain) (LLL/H), as in:

/Nbò/ 'horse' + /mɛ́Ntʂʰɛ́/ 'tail' > /Nbò mèNtʂʰɛ́/ 'horse tail'
/ŋwæ̀/ 'cow' + /Ndʐɤ̀pí/ 'skin, leather' > /ŋwæ̀ Ndʐɤ̀pí/ 'cowhide'

(iii) The third, less frequent, tonal pattern is that in which the domain of tone change appears to be restricted to the first two syllables, whereas the remaining syllable(s) is realized with a low tone (HHL), as in:

/ṣɛ́/ 'iron' + /Ndʐɤ́-Ndʐɤ́/ 'button' > /ṣɛ́ Ndʐɤ́-Ndʐɤ̀/ 'iron button'
/əmɛ́/ 'soldier' + /Ndʐòmó/ 'official' > /əmɛ́ Ndʐómò/ 'general'
/ɜ̃́-ṣɤ́/ 'the Ersu people' + /xó/ 'language' > /ɜ̃́-ṣɤ́ xò/ 'the Ersu language'

Ersu function words and discourse particles (e.g. the genitive particle /i/, the focus particle /ŋɛ/ in the recorded text) are never pronounced in isolation. Their surface tone realization depends on the tone of the preceding (host) lexical word (as in tonal contours in compounds).

Transcription of the recorded text: 'The North Wind and the Sun'
The original audio and video recordings (made with a Digidesign 003 Rack soundcard, Pro Tools LE software for iMac, an AKG C520L headset microphone, and video cameras Sony HDR-XR 520E and Sony HDR-PJ650) have been made available to the *JIPA* along with this analysis.

Semi-narrow phonetic transcription
tɛ́ njó tɛ́ kɛ́ nɛ́ | mɛ́ɜ̃ lá mátʂʰá = dʐí dè-dʐɤ̀-dʐɤ̀-á dʐì ‖ nè nɛ́ = wó tʰɛ́ tá-wá xíbá ká | xíbá ká xá nɛ́ nɛ́ = wó dè-Ndʐí-Ndʐí | ṣétɛ́ = wó xá nɛ́ ṣétɛ́ = wó já-ŋá dʐì-á | ṣétɛ́ = wó xá nɛ́ ṣétɛ́ = wó sòmó jà-dʐo dʐì-á | Ndʐí-Ndʐí ‖ nɛ́-Ndʐí-Ndʐí nɛ́-Ndʐí-Ndʐí nɛ́ tʰɛ̀-á tɛ́ = wó tɛ́-á tʰà-kʰá mà = pʰá | tʰɛ́ = kɛ́ nɛ́ Ndʐá ɹɤ́-kʰwá Ndʐɤ̀ʂ̀-ṣɤ́ tɛ́ ŋà-lá dʐì-gè-á | Ndʐɤ̀ʂ̀-ṣɤ́ lɛ́ ətɕí Ngámɛ́ já-tʂʰá jà-kʰwá tá-tʂʰá dè-zɤ́ tʂá dʐì-gɛ́ ‖ nè nɛ́ = wó tʰɛ́ ná-tɕá-jí | jó = dʐí ṣétɛ́ = wó já-ŋá nɛ́ tʰɛ́ Ndʐɤ̀ʂ̀-ṣɤ́ Ngámɛ́ jà-kʰwà = tʂʰà | tʰà-kwá ná-lá ṣɤ́ tʰɛ́-já dʐì-á ‖ nɛ́ mɛ́ɜ̃ = wó | jò só tɕʰó-gɛ́ tʰɛ́-já dʐì-gè-á ‖ mɛ́ɜ̃ = wó dè-dè-mæ̀ nɛ́ | átʰɛ́ Ndʐɤ̀ʂ̀-ṣɤ́ Ngámɛ́ = tʂʰá tʰɛ́-tʰɛ́-pɤ́-pɤ́ | mɛ́ɜ̃ = wó dè-dè-ɜ̃̀ nɛ́ | Ngámɛ́ = tʂʰá tʰɛ́-tʰɛ́-pɤ́-pɤ́ | áNdʐí tɕʰò-á xá lá Ngámɛ́ = tʂʰá tʰà-kwà-jí ná-lá mà = pʰá ‖ nɛ́ tʰɛ́ = kɛ́ nɛ́ mátʂʰá = wó tɕʰó = gɛ́ tʰɛ́-já dʐì-á ‖ má tʂʰá = wó tɛ́ kʰó ŋà-sá xá nɛ́ | átʰɛ́ ṣɤ́ dà-tʂʰà-jí dà-là nɛ́ | Ngámɛ́ = tʂʰá dʐógwá tʰà-kwà-jí ná-lá dʐì-gɛ́ ‖ nɛ́ tʰɛ́ = kɛ́ nɛ́ mɛ́ɜ̃ = wó èj | á ná kʰɛ̀-ʂ̀ mà-tó-á tʰɛ́-já dʐì-á ‖

Interlinear morphemic glossing

Abbreviations used in the gloss below follow the Leipzig Glossing Rules (LGR, http://www.eva.mpg.de/lingua/resources/glossing-rules.php). Non-standard abbreviations (those not included in the LGR) are: ANM = animate, CMPR = comparative.

té	njó	té	ké	né |	mɛ̂ɕ	lá	mátʂʰá = d̠zí
one	day	one	LOC	TOP	wind	and	sunshine = DU

dɛ̀-d̠zɤ̀-d̠zɤ̀-á		dzì ‖	nɛ̀	né = wó	tʰɛ́
upward-meet-meet-PFV		say	TOP	two = item	that

tá-wá	xíbá	ká |	xíbá	ká	xá	né	né = wó
one-together	conversation	hit	conversation	hit	time	TOP	two = item

dɛ̀-Ndz̠ì-Ndz̠í	|	ʂétɛ́ = wó	xá	né	ʂétɛ́ = wó
upward-dispute-dispute		who = item	time	TOP	who = item

já-ŋá	dzì-á |	ʂétɛ́ = wó	xá	né	ʂétɛ́ = wó
ITSF-ruthless	say-PFV	who = item	time	TOP	who = item

sòmó	jà-d̠zó	dzì-á |	Ndz̠í-Ndz̠í	‖	né-Ndz̠í-Ndz̠í
strength	ITSF-exist	say-PFV	dispute-dispute		downward-dispute-dispute

né-Ndz̠í-Ndz̠í	né	tʰɛ́-á	tɛ́ = wó	tɛ́-á
downward-dispute-dispute	TOP	that-N-AGT	one = item	one-N-AGT

tʰà-kʰá	mà = pʰá |	tʰɛ́ = kɛ́	né	Ndz̠á	ɻź-kʰwá
outward-win	NEG = be.able	that = LOC	TOP	Chinese	road-big

Ndɻ̩ʐ̀z̠̀-ʂɤ́	té	ŋà-lá	dzì-gè-á |	Ndɻ̩ʐ̀z̠̀-ʂɤ́	lɛ́
walk-NMLZ.AGT	one	outward-come	say-PROG-PFV	walk-NMLZ.AGT	FOC

ətɕí	Ngámɛ́	já-tʂʰá	jà-kʰwá	tá-tʂʰá	dɛ̀-z̠ź
really	clothes	ITSF-set	ITSF-big	one-set	upward-wear

tʂá	dzì-gɛ́ ‖	nɛ̀	né = wó	tʰɛ́	ná-tɕá-jí |	jó = d̠zí
DUR	say-PROG	you	you = item	that	downward-agree-NMLZ	self = DU

ʂétɛ́ = wó	já-ŋá	né	tʰɛ́	Ndɻ̩ʐ̀z̠̀-ʂɤ́	Ngámɛ́
who = item	ITSF-ruthless	TOP	that	walk-NMLZ.AGT	clothes

jà-kʰwà = tʂʰà	|	tʰà-kwá	ná-lá	ʂɤ́	tʰɛ́-já
ITSF-big = set		here-take.off	downward-come	CAUS	that-kind

dzì-á	‖	né	mɛ̂ɕ = wó |	jò	só	tɕʰó-gɛ́	tʰɛ́-já
say-PFV		TOP	wind = item	self	first	perform-PROG	that-kind

dzì-gè-á	‖	mɛ̂ɕ = wó	dè-dè-má	né |	átʰɛ́
say-PROG-PFV		wind = item	upward-upward-blow	TOP	that

Ndɻ̩ʐ̀z̠̀-ʂɤ́	Ngámɛ́ = tʂʰá	tʰɛ́-tʰɛ́-pɤ́-pɤ́ |	mɛ̂ɕ = wó
walk-NMLZ.AGT	clothes = set	here-here-wrap-wrap	wind = item

dè-dè-ɚ̀	né |	Ngámɛ́ = tʂʰá	tʰɛ́-tʰɛ́-pɤ́-pɤ́ |
upward-upward-blow	TOP	clothes = set	here-here-wrap-wrap

áNdz̠í	tɕʰó-á	xá	lá	Ngámɛ́ = tʂʰá	tʰà-kwà-jí
how	perform-PFV	time	and	clothes = set	here-take.off-NMLZ

ná-lá	mà = pʰá	‖	nɛ̀	tʰɛ́ = kɛ́	né
downward-come	NEG = be.able		TOP	that = LOC	TOP

mátʂʰá = wó   tɕʰó-gɛ́   tʰɛ́-já   dʐì-á || mátʂʰá = wó
sunshine = item   perform-PROG   that-kind   say-PFV   sunshine = item

tɛ́   kʰó   ŋà-ʂá   xá   nɛ́| átʰɛ́   ʂv́   dà-tʂʰà-jí
one   shine   outward-put   time   TOP   that   person   upward-hot-NMLZ

dà-là   nɛ́ | Ngámɛ́ = tʂʰá   dʐógwá   tʰà-kwà-jí
upward-come   TOP   clothes = set   immediately   here-take.off-NMLZ

ná-lá   dʐì-gɛ́ || nɛ́   tʰɛ́ = kɛ   nɛ   mɛ́ɚ̃ = wó   èj |
downward-come   say-PROG   TOP   that = LOC   TOP   wind = item   ITRJ

á   ná   kʰɛ̀-ʂź̩   má = tó-á   tʰɛ́-já   dʐì-á ||
I   you.N-AGT   inward-defeat   NEG = can-PFV   that-kind   say-PFV

Acknowledgements

References

Anderson, Victoria B. 2008. Static palatography for language fieldwork. *Language Documentation and Conservation* 2(1), 1–27.

Bradley, David. 1979. *Proto-Loloish* (Scandinavian Institute of Asian Studies, monograph series, 39). London & Malmo: Curzon Press.

Bradley, David. 1985. Nasality in Bisu and Bisoid. In Suriya Ratanakul, David Thomas & Suwilai Premsrirat (eds.), *Southeast Asian Linguistic Studies Presented to André-G. Haudricourt*, 234–263. Bangkok: Institute of Language and Culture for Rural Development, Mahidol University.

Bradley, David. 1997. Tibeto-Burman languages and classification. In David Bradley (ed.), *Tibeto-Burman languages of the Himalayas* (Papers in Southeast Asian Linguistics 14), 1–71. Canberra: Department of Linguistics, Research School of Pacific and Asian Studies, Australian National University.

Bradley, David. 2008. Labiovelars in Adur. Presented at the Workshop on Tibeto-Burman Languages of Sichuan, Academia Sinica, Taipei, 21–22 November 2008.

Chao, Yuen Ren. 1972. *Mandarin primer: An intensive course in spoken Chinese*. Cambridge, MA: Harvard University Press & London: Oxford University Press. [First published in 1948 by Harvard University Press.]

Chirkova, Katia. 2012. The Qiangic subgroup from an areal perspective: A case study of languages of Muli. *Language and Linguistics* 13(1), 133–170.

Chirkova, Katia. 2014. The Duoxu language and the Ersu–Lizu–Duoxu relationship. *Linguistics of the Tibeto-Burman Area* 37(1), 104–146.

Chirkova, Katia & Yiya Chen. 2013. Lizu. *Journal of the International Phonetic Association* 43(1), 75–86.

Chirkova, Katia & Zev Handel. 2013a. Diachronic developments of voiceless nasals in Ersu, Lizu, and related languages. Presented at the 46th International Conference on Sino-Tibetan Languages and Linguistics, Dartmouth College, New Hampshire, August 7–10. [Written version of the presentation is available at https://www.academia.edu/7637522/, accessed July 2014.]

Chirkova, Katia & Zev Handel. 2013b. Fricativization of high vowels and glides in Ersu, Lizu, and related languages. Presented at the 3rd Workshop on Sino-Tibetan Languages of Sichuan, Paris, 2–4 September 2013.

Ladefoged, Peter. 2003. *Phonetic data analysis: An introduction to fieldwork and instrumental techniques*. Malden, MA: Blackwell.

Ladefoged, Peter & Ian Maddieson. 1990. Vowels of the world's languages. *Journal of Phonetics* 18, 93–122.

Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. Malden, MA: Blackwell.

Lama Ziwo. 1994. Namuyiyu zhishu yanjiu [On the genetic affiliation of the Namuyi language]. *Minzu Yuwen* [Minority languages] 1, 50–60.

Laver, John. 1994. *Principles of phonetics*. Cambridge: Cambridge University Press.

Li Lan. 2010. Sichuan Muli xian fangyan jilue [The Chinese dialect of Muli, Sichuan: A survey]. *Fangyan* [Dialects] 2, 114–133.

Li Min & Ma Ming. 1983. *Liangshan yiyu yuyin gailun* [Introduction to the phonetics of the Yi Language of Liangshan Prefecture]. Chengdu: Sichuan Nationalities Press.

Liu Huiqiang. 2007. Ersuyu gaikuang [An outline of the Ersu language]. In Li Shaoming & Liu Junbo (eds.), *Ersu Zangzu yanjiu* [Studies on the Ersu Tibetans], 462–500. Beijing: Minzu Chubanshe. [First published in 1983 in Sichuan sheng Minzu Yanjiusuo (eds.), *Minzu yanjiu lunwenji* [Collected papers on minority languages], vol. 1.]

Marchal, Alain. 1988. *La palatographie*. Paris: Editions du Centre National de la Recherche Scientifique.

Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman: System and philosophy of Sino-Tibetan reconstruction*. Berkeley, CA: University of California Press.

Matisoff, James A. 2006. Much Adu about something: Extrusional labiovelars in a Northern Yi patois. *Linguistics of the Tibeto-Burman Area* 29(1), 95–106.

McGowan, R. S. 1992. Tongue-tip trills and vocal-tract wall compliance. *Journal of the Acoustical Society of America* 91, 2903–2910.

Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In Peter F. MacNeilage (ed.), *The production of speech*, 189–216. New York: Springer.

Pan, Zhengyun. 2001. Yiyu Aduhua chunruan'e fufuyin shengmu bijiao yanjiu [A comparative study of labiovelar cluster initials in the Adu patois of the Yi language]. *Minzu Yuwen* [Minority languages] 2, 17–22.

Schmidt, Edwin. 2011. Commodification in an Ersu Tibetan village of Sichuan, China. MA dissertation, Oregon State University.

Spajić, Siniša, Peter Ladefoged & P. Bhaskararao. 1996. The trills of Toda. *Journal of the International Phonetic Association* 26(1), 1–21.

Sun, Hongkai. 1982. Ersu (Duoxu) hua jianjie [An outline of Ersu (Duoxu)]. *Yuyan Yanjiu* [Linguistic study] 2, 241–264.

Sun, Hongkai. 1983. Liujiang liuyu de minzu yuyan ji qi xishu fenlei [Minority languages of the Six River Valley and their genetic classification]. *Minzu Xuebao* [Scholarly Journal of Nationalities] 3, 99–274.

Sun, Hongkai. 2001. Lun Zang-Mian yuzu zhong de Qiang yuzhi yuyan [On the Qiangic branch of the Tibeto-Burman language family]. *Language and Linguistics* 2(1), 157–181.

Sun, Hongkai. 2009. The Ersu Shaba pictographic writing system. *Journal Asian Highlands Perspectives* 1, 159–186.

Sun, Jackson T.-S. 1997. The typology of tone in Tibetan. *Chinese languages and linguistics IV: Typological studies of languages in China* (Symposium Series of the Institute of History and Philology – Academia Sinica), 485–521. Taipei: Taiwan.

Wang, Dehe. 2010. *Ersu zangzu wenhua yanjiu* [Study of Ersu Tibetan culture]. Chengdu: Sichuan University Press.

Wang, Dehe. 2011. Ersu shaba xiangxingwen lishu yanjiu [Divination books and pictographic script of the Ersu Shaba priests]. *Cahiers de linguistique – Asie orientale* 40(2), 225–248.

Wu Da. 2010. *Zuqunxing yu zuqun rentong jiangou – Sichuan Ersuren de minzuzhi yanjiu* [The construction of ethnic identities among the Ersu of Sichuan]. Beijing: Minzu Chubanshe.

Yu, Dominic. 2012. *Proto-Ersuic*. Ph.D. dissertation, University of California, Berkeley.

Zhang, Sihong. 2013. *A reference grammar of Ersu: A Tibeto-Burman language of China*. Ph.D. dissertation,
James Cook University.

Zhang, Sihong. 2014. The expression of knowledge in Ersu. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds.), *The grammar of knowledge: A cross-linguistic typology*, 132–147. Oxford: Oxford University Press.