DEPARTMENT OF LINGUISTICS AND TRANSLATION
LT4235 PROJECT

An Acoustic Analysis of Mandarin Sibilants Produced By Cantonese Speakers

Supervisor: Dr. Lee Wai Sum, Vanti

Undergraduate student: Wong Cheuk Kin

Student number: (removed for privacy)

Major: Linguistics and Language Technology

Date: May, 2015

Abstract

This study investigates the production of the three sets of sibilants in Mandarin, i.e. the denti-alveolar [ʦ, ʦʰ, s], post-alveolar (or retroflex) [tʂ, tʂʰ, ʂ], and alveolo-palatal [ʨ, ʨʰ, ɕ], produced by four university students, 2 male and 2 female, in Hong Kong whose native language is Cantonese, through acoustic analysis. Frequency values for the noise peak and noise range, which are two major acoustic properties for distinction among the sibilants in different place categories, were measured for the test sibilants produced by the subjects. In comparison of the data from a native speaker of Mandarin, some patterns of the sibilant production are generalized for the Cantonese subjects.

Generally speaking, the Cantonese subjects have not mastered the distinction of the three sets of Mandarin sibilants. The denti-alveolar sibilants [ʦ, ʦʰ, s] are most frequently mispronounced by the subjects; both denti-alveolar [ʦ, ʦʰ, s] and alveolo-palatal [ʨ, ʨʰ, ɕ] sibilants cannot be identified as one place category in most cases, which are not produced within a boundary of anyone of the three sets of sibilants and/or not consistently mixed up with other sibilant equivalents. The production of the retroflex [tʂ, tʂʰ, ʂ] is the best, and they can be clearly differentiated from the other two sets of sibilants. The patterns of errors in the production of Mandarin sibilants for the Cantonese subjects are similar to those for the Korean, Japanese, and Vietnamese subjects reported in Chung and Si (2009). The findings of this study contributed to our deeper understanding towards the difficulties in the production of Mandarin sibilants of Cantonese learners, paving the way for Mandarin teaching or further investigation.

Keywords: acoustic analysis, Mandarin sibilants, noise peak and noise range, L2 Mandarin learning

Table of Content

1. Introduction and background information

English and Cantonese are two official languages used in Hong Kong. Since the Handover of Hong Kong to China in 1997, the promotion and usage of Mandarin, the official language and lingua franca in China, have been increasing. Under the implementation of the language policy ‘Bi-literacy and Tri-lingualism’ in Hong Kong, two Chinese varieties, Cantonese and Mandarin, are now taught in most of the local schools at all the kindergarten, primary, secondary, and tertiary levels. As a result, more and more people in Hong Kong can speak both Cantonese and Mandarin, although Cantonese is the dominant language as well as the first language (L1) of most Hong Kong people and Mandarin as their second language (L2).

Cantonese and Mandarin are conventionally considered as two dialects of the Chinese language, but they indeed differ a lot in the sound system as well as vocabulary and grammar. This leads to a great difficulty for L1 Cantonese speakers in L2 Mandarin learning. A comparison of the sound systems of Cantonese and Mandarin shows that there is an apparent difference in the sibilant consonants, including the coronal fricatives and affricates, between the two Chinese varieites.

According to Lee and Zee (2003), there are three sets of coronal sibilants of different place categories in Mandarin, namely denti-alveolar [ʦ, ʦ^{h}, s], retroflex or apical post-alveolar [tʂ, tʂ^{h}, ʂ], and alveolo-palatal [ʨ, ʨ^{h}, ɕ], as presented in Table 1. In Cantonese, according to Zee (1999), there is only one set of alveolar sibilants [ʦ, ʦ^{h}, s] as presented in Table 2.

Place / Sibilant types	Dental	Alveolar	Post-alveolar	Palatal
Affricate	ʦ ʦ^{h}		tʂ tʂ^{h}	ʨ ʨ^{h}
Fricative	s		ʂ	ɕ

Table 1. Denti-alveolar, retroflex, and alveolo-palatal affricate and fricative sibilants in Mandarin.

Place / Sibilant types	Alveolar
Affricate	ʦ ʦʰ
Fricative	s

Table 2. Alveolar affricate and fricative sibilants in Cantonese.

All the coronal affricates and fricatives in Mandarin and Chinese are voiceless and produced with high-pitch hissing noise which is a prominent acoustic feature for sibilant sounds (Ladefoged and Johnson, 2011).

Among the three sets of sibilants in Mandarin, the denti-alveolar [ʦ, ʦʰ, s] are similar to the alveolar [ʦ, ʦʰ, s] in Cantonese, but the contact between the tongue and the alveolar ridge extends forward to the dental area in Mandarin [ʦ, ʦʰ, s] (Lee and Zee, 2003), while it extends backward to the post-alveolar area in the Cantonese [ʦ, ʦʰ, s] (Zee, 1999). The other two sets of sibilants in Mandarin are the apical post-alveolar [tʂ, tʂʰ, ʂ], which are often referred to as retroflexes, degree of tip curling, and alveolo-palatal [ʨ, ʨʰ, ɕ], which are made with extensive contact between the front part of the tongue, including the tongue blade and front dorsum, and the alveolar and post-alveolar areas (Lee and Zee, 2003).

Apart from the difference in place category of the sibilants, Cantonese and Mandarin sibilants also differs in the adjacent vowel when they occur in open CV syllables. In both Mandarin and Cantonese, the sibilant consonants only occur in word- or syllable-initial position. In CV syllables as presented in Table 3, the Mandarin [ʦ, ʦʰ, s] are followed by [ɿ, a, u, ɤ], whereas the equivalents [tʂ, tʂʰ, ʂ] followed by [ʅ, a, u, ɤ] and [ʨ, ʨʰ, ɕ] followed by [i, y]. It follows that [ʦ, ʦʰ, s] and [tʂ, tʂʰ, ʂ] are in contrastive distribution, as they can be followed by the same type of vowel in CV syllables, whereas [ʨ, ʨʰ, ɕ] are in complementary distribution with [ʦ, ʦʰ, s] and [tʂ, tʂʰ, ʂ], as the formers are followed by a high front or palatal vowel only. According to Lee and Zee (2003), the Mandarin [ɿ] and [ʅ] represented with the non-IPA symbols are often referred to as the apical vowels by the Chinese linguists. The two apical vowels are considered as the variants of the vowel [i] when preceded by the sibilant consonants. [ɿ] is produced as apical denti-alveolar when the preceding sibilants are the denti-alveolar [ʦ, ʦʰ, s], whereas [ʅ] is produced as apical retroflex when the preceding sibilants are the retroflex [tʂ, tʂʰ, ʂ] (Lee and Zee, 2003).

	Syllable-initial sibilants	Position in CV syllables
Mandarin	[ʦ, ʦʰ, s]	'__ [ɿ, a, u, ɤ]'
	[tʂ, tʂʰ, ʂ]	'__ [ʅ, a, u, ɤ]'
	[ʨ, ʨʰ, ɕ]	'__ [i, y]'
Cantonese	[ʦ, ʦʰ, s]	'__ [i, y, ɛ, a, ɔ]'

Table 3. Initial positions of the Mandarin and Cantonese sibilant consonants in CV syllables.

What has been described about the variations of the vowel [i] and the occurrence of the two apical vowels [ɿ] and [ʅ] are not in Cantonese. In Cantonese, the single set of alveolar sibilants [ʦ, ʦʰ, s] can be followed by any one of the five vowels [i, y, ɛ, a, ɔ], including two high front vowels [i, y] (Zee, 1999). Therefore, the Cantonese sibilants [ʦ, ʦʰ, s] are similar to the Mandarin alveolo-palatal sibilants [ʨ, ʨʰ, ɕ] in terms of the adjacent vowels.

Due to the differences in the type of the sibilants and the type of vowel adjacent to the sibilants in CV sibilants between Mandarin and Cantonese, it is expected that the production of the Mandarin sibilant consonants is difficult for Cantonese speakers. Some difficulties have been reported in a number of previous studies.

Tsang (1996), based on his observation in Mandarin teaching to Cantonese-speaking students, considered that Cantonese learners are easily confusing the three sets of sibilants in Mandarin. According to the author, while the retroflex or apical post-alveolar sibilants [tʂ, tʂʰ, ʂ] do not occur in Cantonese, they are not the most difficult ones for Cantonese speakers, as they can be substituted with the similar post-alveolar sibilants [ʤ, ʧ, ʃ] borrowing from English, another L2 of Cantonese speakers. As for the Mandarin alveolo-palatal [ʨ, ʨʰ, ɕ], they are similar to the Cantonese alveolar sibilants [ʦ, ʦʰ, s] articulatory with contact extensively made on the whole alveolar ridge, so they are not a problem as well. The Mandarin dent-alveolar sibilants [ʦ, ʦʰ, s] are the most difficult ones to be produced by Cantonese speakers, especially when they are followed by the apical vowel [ɿ]. The sibilants [ʦ, ʦʰ, s] are commonly mispronounced as the alveolo-palatal equivalents [ʨ, ʨʰ, ɕ], such as in [sɿ] 四 'four' which becomes [ɕi] 戲 'film'.

Ng (2001), a contrastive study of the sound systems of Cantonese and Mandarin, has a view different from Tsang (1996) and believed that the Mandarin retroflex sibilants [tʂ, tʂʰ, ʂ], which do not occur in Cantonese, are relatively more difficult to be learnt by Cantonese speakers than the other two types of Mandarin sibilants. Based on her observation, some Cantonese-speaking learners of Mandarin may only pay attention to curl the tongue tip, without curling the two sides of the tongue as well during the production of the Mandarin retroflex sibilants. But, she pointed out that as the Cantonese alveolar sibilants [ʦ, ʦʰ, s] are similar to the Mandarin alveolo-palatal sibilants [ʨ, ʨʰ, ɕ] in actual pronunciation, the Cantonese learners tend to use the Cantonese sibilants [ʦ, ʦʰ, s] to replace the Mandarin alveolo-palatal equivalents, which is the negative transfer from L1 Cantonese to L2 Mandarin.

Lee-Wong (2013) thought that the pronunciation errors in Mandarin made by Cantonese-speaking learners are not merely resulted from the negative transfer from Cantonese, but also due to hypercorrection to certain types of sounds. She pointed out the misunderstanding of many Cantonese speakers that there are a lot of retroflex sounds in Mandarin, leading them in using retroflex sibilants to substitute with other sibilant equivalents. This case of hypercorrection indicates that Cantonese speakers are unable to distinguish the different types of sibilants in Mandarin. Therefore, the author suggested that Mandarin teachers should not only pay attention to the possible difficulties in producing the different sibilants, but also to enhance the ability of students to differentiate among the sibilants.

Li (2009), based on the observation of the errors made by 48 Cantonese-speaking university students in one-year Mandarin lessons, generalized that the students easily mispronounced among the three types of fricative sibilants [s, ʂ, ɕ] and also among the three types of unaspirated affricate sibilants [ʦ, tʂ, ʨ]. The author concluded that the errors in Mandarin pronunciation are similar among the Cantonese students in the class, which are mainly resulted from the negative transfer from their L1.

In Hon (2003), audio recordings of a word list consisting of the three sets of Mandarin sibilants produced by 30 Cantonese speakers, 20 university students and 10 secondary students, were made. Based on the author's perceptual judgment of the recordings, mispronunciations of the Mandarin sibilants were mainly observed in the speech of about 60% of the subjects, and 4 most frequent cases of errors were found: 1) confusing the three unaspirated affricates [ʦ, tʂ, ʨ] among each other, 2) the retroflex fricative [ʂ] → the alveolo-palatal fricative [ɕ], 3) the aspirated retroflex affricate [tʂʰ]→ the unaspirated alveolo-palatal affricate [ʨ], 4) the aspirated alveolo-palatal affricate [ʨʰ] → the aspirated denti-alveolar affricate [ʦʰ]. Of these four cases, the first one is the most common across the subjects. Furthermore, the subjects also tend to substitute the Mandarin retroflex [tʂ, tʂʰ, ʂ] and denti-alveolar [ʦ, ʦʰ, s] sibilants with the Cantonese alveolar [ʦ, ʦʰ, s], such as [ʂʅ] 師 'teacher' and [sɿ] 司 'in charge of' → [si]; [tʂʅ] 知 'know' and [ʦɿ] 資 'capital' → [ʦi]; [tʂʰʅ] 持 'hold' and [ʦʰi] 磁 'magnet' → [ʦʰi], demonstrating the negative transfer from L1 Cantonese to L2 Mandarin.

Similar to Hon (2003), Wu and Su (2014) examined the pronunciation errors made in the audio recordings of a list of Mandarin words produced by 7 female Cantonese-speaking subjects who have been studying in Taiwan for one to three years. The authors also evaluated the Speech learning Model proposed by Flege (1995) on the prediction of the errors of the Cantonese subjects. This model classifies the sounds in L2 or the target language into three types by comparison with the sounds in L1 or the native language: 1) identical phones, 2) similar phones, and 3) new phones. Identical phones refer to the same sounds in both the native and target language; similar phones are the sounds which are similar, but not identical, in quality between the native and target languages; new phones are the sounds of the target language which do not exist in the native language. Based on the model, the sibilants [ʦ, ʦʰ, s] which are identical phones in Cantonese and Mandarin, are predicted to be produced with no errors, and [tʂ, tʂʰ, ʂ] and [ʨ, ʨʰ, ɕ],which are new phones, are easily to be learnt and also produced with no errors. The speech data from the Cantonese subjects in Wu and Su (2014) show that for the Mandarin fricatives and affricates, no errors are found in all sibilants [ʦʰ, s, ʂ, ʨ, ɕ]. As for retroflex sibilants, only [ʂ] is correctly pronounced and errors are observed in the other two retroflex sibilants [tʂ, tʂ]. Generally speaking, the model can predict the pronunciation of the Mandarin sibilants by Cantonese speakers, except for the retroflex sibilants. The findings suggest that some of the Mandarin sibilants are still difficult sounds for Cantonese speakers, even they have lived in a Mandarin-speaking place for over 1 year.

What presented before about the errors in Mandarin pronunciation for Cantonese speakers in the previous studies (Tsang, 1996; Ng, 2001; Lee-Wong, 2013; Li, 2009; Hon, 2003) are mainly based on the authors' impressionistic or perceptual judgment, without providing any measured acoustic data for substantiation.

Chung and Si (2009) conducted an acoustic analysis of the Mandarin fricative sibilants produced by three groups of L2 learners of Mandarin, Japanese, Korean, and Vietnamese. Speech data of each group were provided by 2 male subjects aged from 21-25, who have learnt Mandarin for about 6-7 months. Two major acoustic features of the noise of the fricatives, namely the peak of the noise and the range of the noise energy distribution (Heiz and Stevens, 1961; Behrens and Blumstein, 1988; Evers et al., 1998; Pawel, 2006), were analyzed and measured. For comparison purpose, the authors also analyzed the speech data from a Mandarin teacher who received Grade 1A in the National Putonghua Proficiency Test.

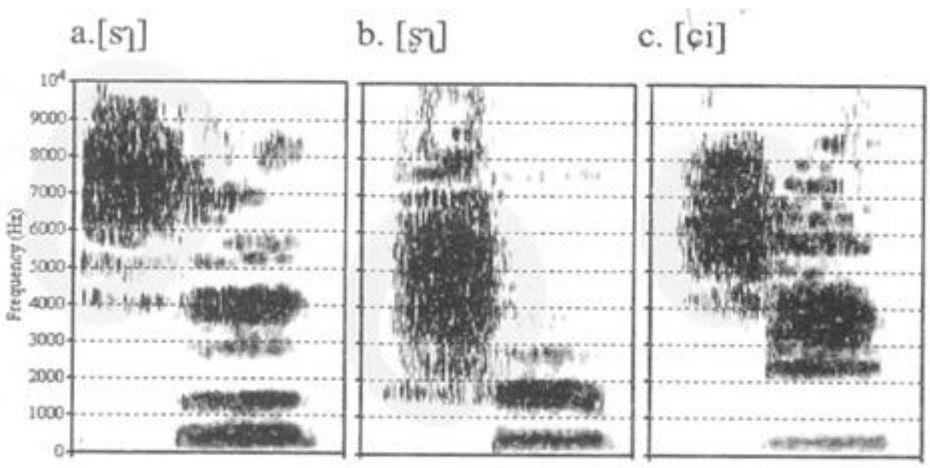


Figure 1. Wide-band spectrograms of the words [sɿ], [ʂʅ], [ɕi] in Mandarin for a native speaker (from Chung and Si, 2009)

Figure 1 from Chung and Si (2009) shows the wide-band spectrograms of the words [sɿ], [ʂʅ], [ɕi] produced by the Mandarin teacher. As can be seen, the noise energy of the word-initial fricatives [s, ʂ, ɕ] concentrates in different frequency regions. The intense noise distributes in the region approximately from 6,000 Hz to 9,500 Hz for [s], from 3,000 Hz to 8,000 Hz for [ʂ], and from 5,000 Hz to 8,500 Hz for [ɕ]. Thus, while there is an overlap in the frequency range of the noise distribution among the three fricatives, there is a noticeable difference in the lower boundary of the frequency range of the noise, which is the lowest for [ʂ], followed by [ɕ] and [s] in increasing order.

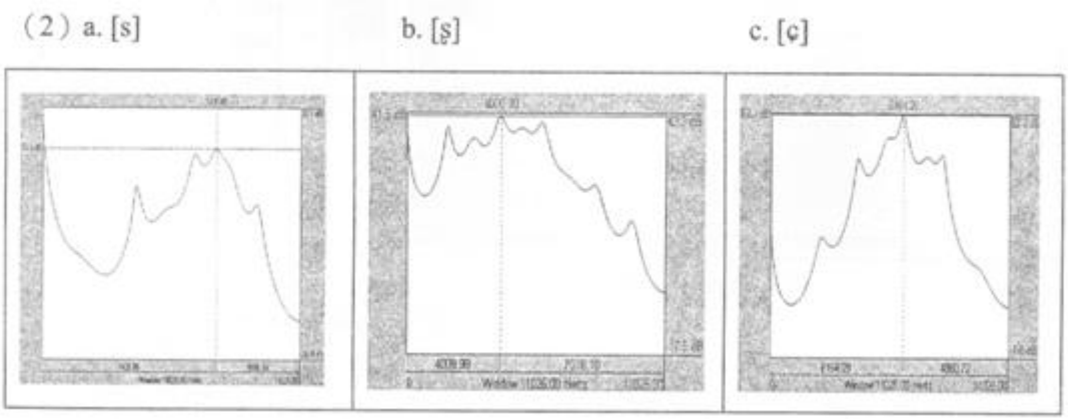


Figure 2. LPC spectra of the noise peaks of the fricative [s], [ʂ], [ɕ] for a Mandarin native speaker (from Chung and Si, 2009)

Figure 2 also from Chung and Si (2009) shows the LPC spectra of the noise peaks measured at a time point of the three fricatives [s, ʂ, ɕ] for the native Mandarin speaker. The measured frequency values for the strongest noise peaks of the three fricatives in descending order are [s] 7428 Hz > [ɕ] 6163 Hz > [ʂ] 4007Hz. This also indicates the highest frequency of noise for [s], the lowest for [ʂ], and medial for [ɕ].

In comparison of the speech data from the native Mandarin speaker, Chung and Si concluded that a correct pronunciation of the Mandarin fricatives [s, ʂ, ɕ] depends on the transfer of the native languages. For the Korean subjects, their pronunciation is better in [s] than in [ʂ, ɕ] because the noise peak of the Korean fortis or tense [s*] is similar to that of the Mandarin [s]. As for the Japanese subjects, they tend to use [s] and [ʃ] of Japanese to replace [s] and [ɕ] of Mandarin, while the Vietnamese subjects tend to use [s] of Vietnamese to replace the three fricatives [s, ʂ, ɕ] in Mandarin. In general, among the three Mandarin fricatives produced by the learners, [ʂ], which does not occur in L1 of the learners, is relatively closer to the target [ʂ] in Mandarin.

In my knowledge, there is a paucity of acoustic studies of the Mandarin sibilant pronunciation by Cantonese speakers in Hong Kong. Also, as pointed out in Cheng (1995) and Zhu, Chen, and Wen (2012), there are not many studies about Mandarin proficiency of Cantonese-speaking students in tertiary institutions as in primary and secondary schools in Hong Kong.

The aim of this study is to investigate the production of the three sets of sibilants in Mandarin, i.e., the denti-alveolar [ʦ, ʦʰ, s], retroflex [tʂ, tʂʰ, ʂ], and alveolo-palatal [ʨ, ʨʰ, ɕ], by Cantonese-speaking university students in Hong Kong through acoustical analysis of their speech. In the light of the fact that there is only one set of alveolar sibilants [ʦ, ʦʰ, s] in Cantonese, the issue of the influence of L1 Cantonese on L2 Mandarin acquisition will be addressed and discussed with the measured frequency data on the noise of the Mandarin sibilants produced by Cantonese subjects. Also, by comparison with the corresponding data from a native Mandarin speaker in this study, the characteristics and patterns of the errors made by Cantonese speakers will be described and generalized.

2. <u>Methodology</u>

2.1 Subjects

In this study, a total of four native Cantonese speakers, 2 males and 2 females, were invited to take part in an individual audio recording. They are aged from 18 to 23, studying undergraduate program in City University of Hong Kong. All of them have taken an elementary Mandarin course, with the knowledge of the sound system and the pronunciation of the sounds of Mandarin. Their proficiency levels in Mandarin are similar, receiving a grade which is above average in the elementary course. For comparison purpose, speech data from a native female speaker of Mandarin were recorded and used as reference. The native speaker is a Mandarin teacher in City University of Hong Kong and is the examiner of the National Putonghua Proficiency Test of Chinese University of Hong Kong.

All the subjects of this study were volunteers. They were told that the purpose of the study is to investigate the Mandarin pronunciation of Cantonese speakers, without informing the focus is on the sibilant consonants of the words to avoid being conscious of the production of the sibilants. For the native Mandarin speaker, she was also told that her speech was used as reference to compare the Mandarin pronunciation made by some Cantonese speakers.

2.2 Test Materials

The test materials used for the investigation were 18 monosyllabic Chinese words of CV structure containing nine target sibilants in Mandarin, including the denti-alveolar [ʦ, ʦʰ, s], retroflex [tʂ, tʂʰ, ʂ], and alveolo-palatal [ʨ, ʨʰ, ɕ] affricates and fricatives. The test words presented in Table 4 show that any one of the target sibilants occurs in the word-initial position and in two different vowel contexts. That is, the denti-alveolar [ʦ, ʦʰ, s] are followed by [ɿ] and [a], the retroflex [tʂ, tʂʰ, ʂ] followed by [ʅ] and [a], and alveolo-palatal [ʨ, ʨʰ, ɕ] followed by [i] and [ia]. As presented earlier in Section 1, in Mandarin the two apical vowels [ɿ] and [ʅ] are considered as the variants of the high front vowel [i], where [ɿ] is restricted to occur after the denti-alveolar sibilants [ʦ, ʦʰ, s] and [ʅ] after the retroflex [tʂ, tʂʰ, ʂ]. As for the alveolo-palatal sibilants [ʨ, ʨʰ, ɕ], they are restricted to occur before a high front vowel, such as the monophthong [i] and the diphthong [ia] beginning with [i]. Since all the selected test words are associated with a high tone [˥], the three sets of target sibilants occur in a similar phonetic context.

Target sibilants	Vowel contexts	Test words		
[ʦ, ʦʰ, s]	[ɿ]	[ʦɿ˥] 資 'capital'	[ʦʰɿ˥] 疵 'flaw'	[sɿ˥] 司 'in charge of'
	[a]	[ʦa˥] 紮 'tie'	[ʦʰa˥] 擦 'rub'	[sa˥] 撒 'cast'
[tʂ, tʂʰ, ʂ]	[ʅ]	[tʂʅ˥] 知 'know'	[tʂʰʅ˥] 吃 'eat'	[ʂʅ˥] 失 'lose'
	[a]	[tʂa˥] 渣 'residue'	[tʂʰa˥] 叉 'folk'	[ʂa˥] 沙 'sand'
[ʨ, ʨʰ, ɕ]	[i]	[ʨi˥] 基 'base'	[ʨʰi˥] 七 'seven'	[ɕi˥] 希 'hope'
	[ia]	[ʨia˥] 加 'add'	[ʨʰia˥] 掐 'nip off'	[ɕia˥] 蝦 'prawn'

Table 4. Test monosyllabic words containing the sibilants [ʦ, ʦʰ, s], [tʂ, tʂʰ, ʂ], and [ʨ, ʨʰ, ɕ] in Mandarin used for investigation.

The selected test words are meaningful in Mandarin, and all Cantonese subjects are familiar with them. The test words presented only in Chinese characters were randomized on a list. Three readings of the word list were given by each subject, making up a total of 270 test tokens (18 test words x 3 repetitions x 5 subjects (4 Cantonese + 1 Mandarin)) for acoustical analysis.

2.3 Audio recordings

Individual audio recordings of the test words from the subjects were performed in the sound-proof booth in the Phonetics Lab at the Department of Linguistics and Translation at City University of Hong Kong. The subjects were instructed to utter the test words consistently at a normal rate of speech and a normal degree of loudness throughout the whole recording. Their speech was recorded using a professional digital recorder with the frequencies up to 44 kHz. The audio signals were subsequently down-sampled to 22 kHz for spectral analysis.

2.4 Data analysis

As presented earlier in Section 1, the primary acoustical feature documented in the literature for distinguishing the different types of sibilants is related to the frequency of the noise energy associated with the sibilants (Heiz and Stevens, 1961; Behrens and Blumstein, 1988; Evers et al., 1998; Nowak, 2006). In this study, the frequencies of the noise peak and the noise energy range are the two acoustical parameters used to compare among the three sets of Mandarin sibilants, the denti-alveolar [ʦ, ʦʰ, s], retroflex [tʂ, tʂʰ, ʂ], and alveolo-palatal [ʨ, ʨʰ, ɕ], produced by the Cantonese subjects, with reference to those from the native speaker. By using the speech analysis software Computerized Speech Lab (CSL 4500) available in the Phonetic Lab, the waveforms showing the sound signals of the test words and the wide-band spectrograms showing the distribution of the noise of the target sibilants of the test words were obtained. At the mid-point of each of the sibilants, Fast Fourier transform (FFT) and Liner predictive coding (LPC) spectral analysis were performed for measurements of the frequencies of the noise range and noise peak of the sibilant.

3. <u>Results</u>

The frequency data on the noise range measured from FFT spectrum and the noise peak measured from LPC spectrum of each of the target sibilants of the test words in Mandarin produced by the four Cantonese speakers and the native speaker are presented on-by-one. The data for the native speaker are presented first, which are used as reference for comparing with those for the Cantonese speakers.

3.1 Native speaker

Table 5 presents the frequency data on the noise peaks of the Mandarin denti-alveolar [ʦ, ʦʰ, s], retroflex [tʂ, tʂʰ, ʂ], and alveolo-palatal [ʨ, ʨʰ, ɕ], by the native Mandarin speaker. Each of the sibilants was produced in two vowel contexts, Context 1 [i]/[ɿ]/[ʅ] and Context 2 [a]/[ia]. The frequency value of the noise peak for any one of the sibilants presented in the table is the mean of three tokens in a particular vowel context.

Target sibilants		Vowel contexts	
		1. [i]/[ɿ]/[ʅ]	2. [a]/[ia]
Denti-alveolar	[ʦ]	7,698	8,470
	[ʦʰ]	7,863	7,587
	[s]	7,441	8,231
Retroflex	[tʂ]	3,270	3,748
	[tʂʰ]	3,160	2,407
	[ʂ]	2,590	2,958
Alveolo-palatal	[ʨ]	7,515	6,467
	[ʨʰ]	6,927	4,629
	[ɕ]	5,879	6,247

Table 5. Frequency values (in Hz) of the noise peaks (mean of three tokens) of the Mandarin sibilants [ʦ, ʦʰ, s], [tʂ, tʂʰ, ʂ], and [ʨ, ʨʰ, ɕ] for a native speaker.

As presented in Table 5, the frequencies of the noise peaks of the three sets sibilants vary in the two vowel contexts, but no general pattern of the difference between the two vowel contexts is observed for the sibilants. For instance, for the denti-alveolar [ʦ, ʦʰ, s], the frequency of the noise peak is higher in Context 2 for [ʦ] and [s], but higher in Context 1 for [ʦʰ]. The pattern is also observed for the retroflex [tʂ, tʂʰ, ʂ], with a higher frequency of the noise peak in Context 2 for [tʂ] and [ʂ], but in Context 1 for [tʂʰ]. As for the alveolo-palatal [ʨ, ʨʰ, ɕ], the frequency of the noise peak is higher in Context 1 for [ʨ] and [ʨʰ], but higher in Context 2 for [ɕ]. Among the three sets of sibilants, the difference in noise peak frequency is apparent. In any vowel context, the frequency of the noise peak is higher for the denti-alveolar [ʦ, ʦʰ, s] (7,441 Hz to 8,470 Hz) than the alveolo-palatal [ʨ, ʨʰ, ɕ] equivalents (4,629 Hz to 7,515 Hz) and much lower for the retroflex [tʂ, tʂʰ, ʂ] equivalents (2,407 Hz to 3,748 Hz). The pattern of the difference in noise peak frequency among the three sets of Mandarin sibilants for the native speaker in this study is similar to that for the native speaker in Chung and Si (2009). The data suggest that the noise peak is an important acoustic parameter for the distinction of the Mandarin sibilants in different place categories.

Table 6 presents the frequency ranges of the noise distribution of the three sets of sibilants [ʦ, ʦʰ, s], [tʂ, tʂʰ, ʂ], and [ʨ, ʨʰ, ɕ] for the native speaker. The frequency range for each sibilant is the mean of three tokens in a particular vowel context. As presented in the table, there is a large overlap in the noise frequency range between the two vowel contexts for each sibilant, indicating no large vowel context effect on the production of the sibilant. Among the three sets of sibilants, while there is also an overlap in the noise frequency range, i.e., 5,483 Hz to 10,610 Hz for [ʦ, ʦʰ, s], 1,139 Hz to 10,067 Hz for [tʂ, tʂʰ, ʂ], and 3,382 Hz to 10,529 Hz for [ʨ, ʨʰ, ɕ], a pronounced difference in the minimum value of the range is observed. Similar to the order of the frequency value for the noise peak among the three sets of sibilants, the descending order of the minimum value of the noise frequency range is denti-alveolar > alveolo-palatal > retroflex. This order was also reported for the native Mandarin speaker in

Chung and Si (2009). Therefore, the minimum value of the noise frequency range is also an important acoustic parameter for the distinction of the Mandarin sibilants in different place categories.

Target sibilants		Vowel contexts	
		1. [i]/[ɿ]/[ʅ]	2. [a]/[ia]
Denti-alveolar	[ʦ]	6,063-10,535	7,560-10,416
	[ʦʰ]	5,483-10,505	6,765-10,488
	[s]	5,970-10,610	6,481-10,513
Retroflex	[tʂ]	1,694-9,118	1,922-9,166
	[tʂʰ]	1,751-10,067	1,139-8,282
	[ʂ]	1,772-9,938	1,987-9,532
Alveolo-palatal	[ʨ]	5,962-10,413	5,540-10,343
	[ʨʰ]	5,605-10,497	3,382-10,529
	[ɕ]	5,118-10,359	3,698-10,440

Table 6. Frequency range (in Hz) of the noise distribution (mean of three tokens) of the Mandarin sibilants [ʦ, ʦʰ, s], [tʂ, tʂʰ, ʂ], and [ʨ, ʨʰ, ɕ] for a native speaker.

Figures 3a to 3c show the frequencies (in Hz) of the noise peak (represented by a vertical bar) and the noise range (represented by an 'I' line) for the sibilants, the unaspirated affricates [ʦ, tʂ, ʨ] (Figure 3a), aspirated affricates [ʦʰ, tʂʰ, ʨʰ] (Figure 3b), and fricatives [s, ʂ, ɕ] (Figure 3c), in the two vowel contexts for the native speaker in this study.

Figure 3a: Unaspirated affricates [ts, tʂ, tɕ]

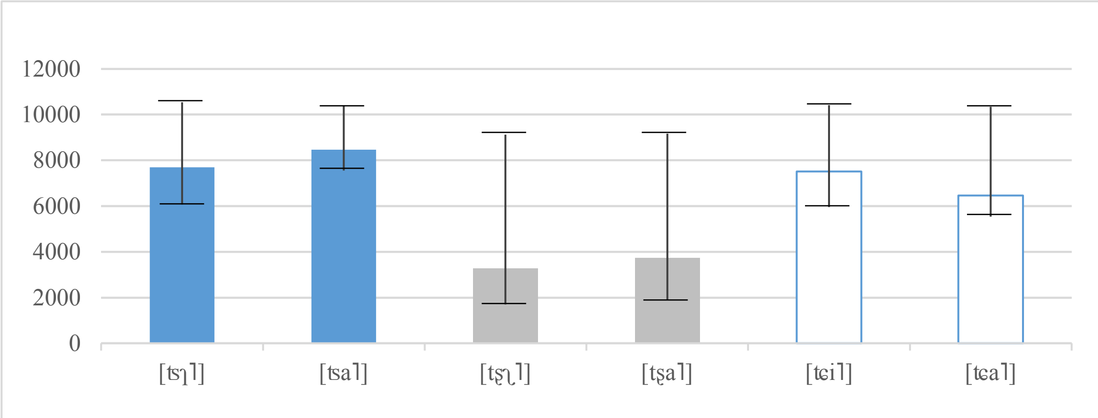


Figure 3b: Aspirated affricates [tsʰ, tʂʰ, tɕʰ]

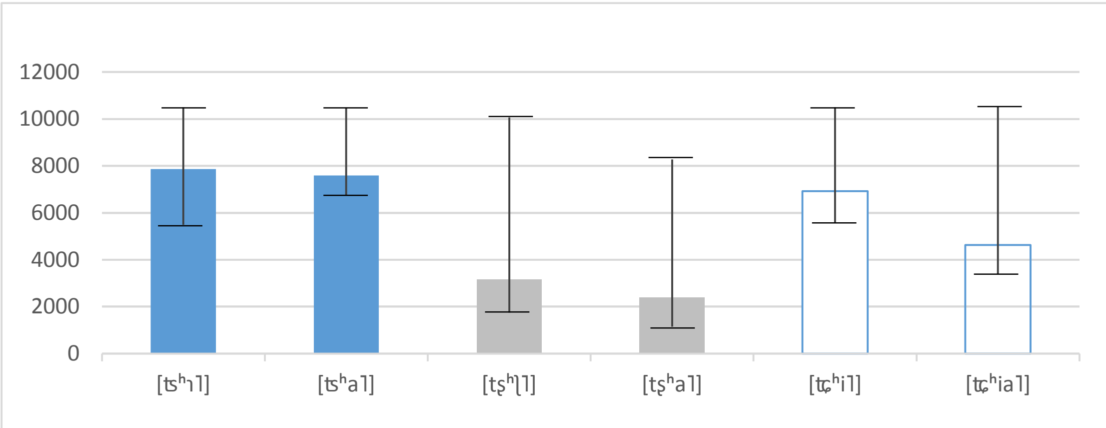


Figure 3c: Fricatives [s, ʂ, ɕ]

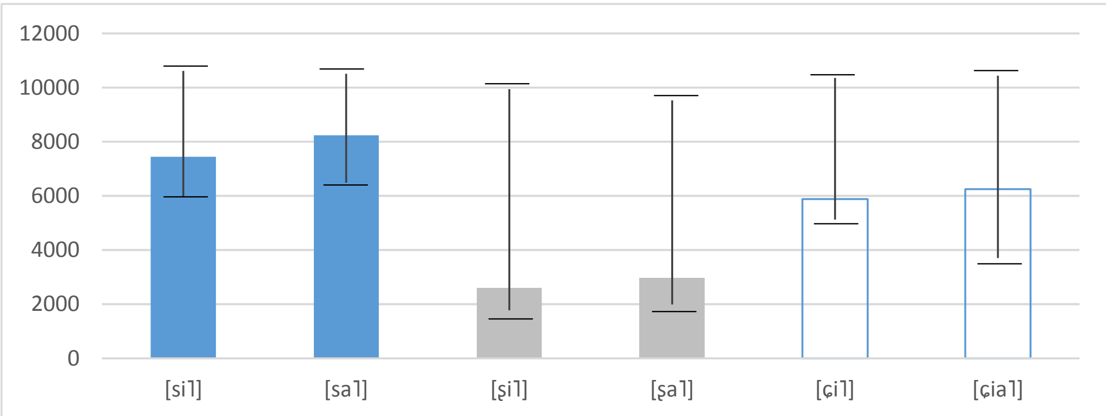


Figure 3a-3c. Frequencies (in Hz) of the noise peak (represented by a vertical bar) and the noise range (represented by an 'I' line) for the unaspirated affricates [ts, tʂ, tɕ], aspirated affricates [tsʰ, tʂʰ, tɕʰ], and fricatives [s, ʂ, ɕ] for a native speaker.

As shown in Figure 3a to 3c for the native speaker, it can be seen that either in terms of the frequency of noise peak or noise range, the difference between the Mandarin sibilants in the two vowel contexts is less pronounced than the difference among the sibilants [ʦ, tʂ, ʨ] (Figure 3a), [ʦʰ, tʂʰ, ʨʰ] (Figure 3b), and [s, ʂ, ɕ] (Figure 3c), in different place categories.

T-test analysis was performed on the differences in the minimum and maximum values of the frequency range of noise between the two vowel contexts for the sibilants of each of the three place categories, the denti-alveolar [ʦ, ʦʰ, s], retroflex [tʂ, tʂʰ, ʂ], and alveolo-palatal [ʨ, ʨʰ, ɕ]. The *p*-value for the differences in each case are presented in Table 7. A *p*-value < 0.05 is marked with an asterisk. As can be seen, there are significant differences in both the minimum and maximum values of the frequency range of noise between the two vowels contexts for [ʦ, ʦʰ, s], but this is not the case for [tʂ, tʂʰ, ʂ], and [ʨ, ʨʰ, ɕ]. For [ʦ, ʦʰ, s] as shown in Figure 3a, the vowel effect on the differences in the two values, especially the minimum value, of the frequency range of noise are not consistent among the three sibilants. Thus, it can be considered that the vowel context effect on the noise distribution of the Mandarin sibilants is not significant.

Sibilants	*P*-values for the difference in the frequency range of noise between the vowel contexts [i/ɿ/ʅ] and [a/ia]	
	Minimum value of the range	Maximum value of the range
[ʦ, ʦʰ, s]	0.005*	0.029*
[tʂ, tʂʰ, ʂ]	0.750	0.040*
[ʨ, ʨʰ, ɕ]	0.005*	0.827

Table 7. *P*-values for the differences in the minimum and maximum values of the frequency range of noise between the two vowel contexts [i/ɿ/ʅ] and [a/ia] for the Mandarin sibilants [ʦ, ʦʰ, s], [tʂ, tʂʰ, ʂ], and [ʨ, ʨʰ, ɕ] for a native speaker. (*$p<0.05$)

Sibilants	Vowel contexts	*P*-values for the difference in the frequency range of noise among the sibilants of the same place category	
		Minimum value of the range	Maximum value of the range
[ʦ, ʦʰ, s]	__ [ɿ]	0.434	0.249
	__ [a]	0.319	0.097
[tʂ, tʂʰ, ʂ]	__ [ʅ]	0.869	0.004*
	__ [a]	0.040*	0.160
[ʨ, ʨʰ, ɕ]	__ [i]	0.351	0.636
	__ [ia]	0.001*	0.117

Table 8. *P*-values for the differences in the minimum and maximum values of the frequency range of noise among the Mandarin sibilants [ʦ, ʦʰ, s], [tʂ, tʂʰ, ʂ], and [ʨ, ʨʰ, ɕ] of the same place category for a native speaker. (*$p<0.05$)

ANOVA analysis was also performed on the differences in the minimum and maximum values of the frequency range of noise among the denti-alveolar [ʦ, ʦʰ, s], retroflex [tʂ, tʂʰ, ʂ], and alveolo-palatal [ʨ, ʨʰ, ɕ] sibilants of the same place categories. The p-values for the differences in each case are presented in Table 8, with $p < 0.05$ marked with an asterisk. As can be seen, there are no cases with a significant difference in any one of the two values of the frequency range of noise among the three denti-alveolar sibilants [ʦ, ʦʰ, s]. For the retroflex tʂ, tʂʰ, ʂ], and alveolo-palatal [ʨ, ʨʰ, ɕ], a significant difference in one of the two values of the frequency range of noise is observed in one or two cases. These data also suggest no significant effect of the manner of articulation of the Mandarin sibilants in each place category on the noise distribution. Therefore, for the native Mandarin speaker in this study, the frequency values for the noise range as well as the noise peak are averaged for all the tokens of the sibilants of the same place category, regardless of the manner of articulation of the sibilants and the vowel context in which the sibilants occur. The data are presented in Table 9, and they are used as reference for comparison of those of the Mandarin sibilants produced by four Cantonese speakers, Cantonese Male 1 and 2, and Female 1 and 2, in this study.

Sibilants	Noise peak	Noise range
Denti-alveolar [ʦ, ʦʰ, s]	7,882	6,387-10,511
Retroflex [tʂ, tʂʰ, ʂ]	3,022	1,711-9,350
Alveolo-palatal [ʨ, ʨʰ, ɕ]	6,277	4,884-10,430

Table 9. Mean frequency values (in Hz) of the noise peak and noise range for the Mandarin sibilants [ʦ, ʦʰ, s], [tʂ, tʂʰ, ʂ], and [ʨ, ʨʰ, ɕ] for a native speaker.

3.2 Cantonese Male 1

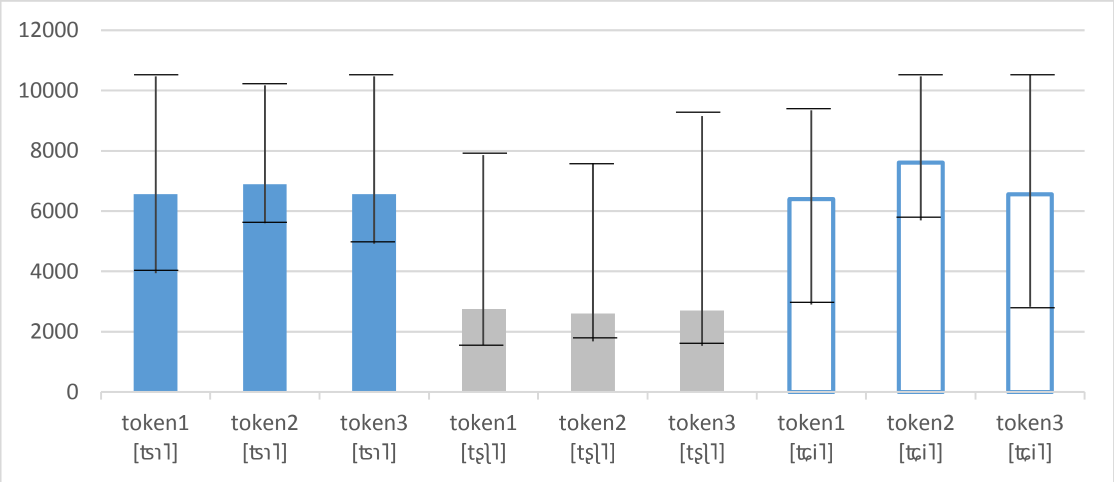


Figure 4. Frequencies in (Hz) of the noise peak (represented by a vertical bar) and noise range (represented by an 'I' line) for the Mandarin unaspirated affricates [ʦ, tʂ, ʨ] in the vowel context [i/ɿ/ʅ] produced by Cantonese Male 1.

Figure 4 shows the frequencies (in Hz) of the noise peak (represented by the vertical bar) and the noise range (represented by an 'I' line) for three tokens of each of the Mandarin unaspirated affricates [ʦ, tʂ, ʨ] in the vowel context [i/ɿ/ʅ] for Cantonese Male 1. Among the three sibilants, the noise peak is noticeably lower for the retroflex [tʂ] (2,590-2,756 Hz) than the denti-alveolar [ʦ] (6,559-6,890 Hz) and the alveolo-palatal [ʨ] (6,394-7,607 Hz). Between [ʦ] and [ʨ], the difference in the noise peak is not pronounced. As for the noise range, a noticeable difference among the three sibilants is in the minimum value of the noise range, which is also the lowest for [tʂ] (1,533-1,679 Hz). The minimum value of the noise range tends to be higher for [ʦ] (3,942-5,597 Hz) than [ʨ] (2,823-2,896 Hz), while it is the highest for one token of [ʨ] (5,695 Hz). By comparison of the noise peak and noise range in a descending order of [ʦ] > [ʨ] > [tʂ] for the native speaker (Figure 3a), it may be considered that Cantonese Male 1 can produce the retroflex [tʂ], but cannot differentiate the denti-alveolar [ʦ] and alveolo-palatal [ʨ].

For Cantonese Male 1, the mean frequencies of three tokens of the noise peak and the minimum value of the noise range are 6,669 Hz and 4,818 Hz for [ʦ], apparently lower than those of the native's [ʦ] (7,882Hz and 6,387 Hz), and 6,853 Hz and 3,804 Hz for [ʨ], similar to those of the native's [ʨ] (6,277 Hz and 4,884 Hz). Thus, it may suggest that Cantonese Male 1 merges the sibilants [ʦ] into [ʨ]. In other words, among the three Mandarin unaspirated affricates [ʦ, tʂ, ʨ], [ʦ] is the most difficult one for Cantonese Male 1.

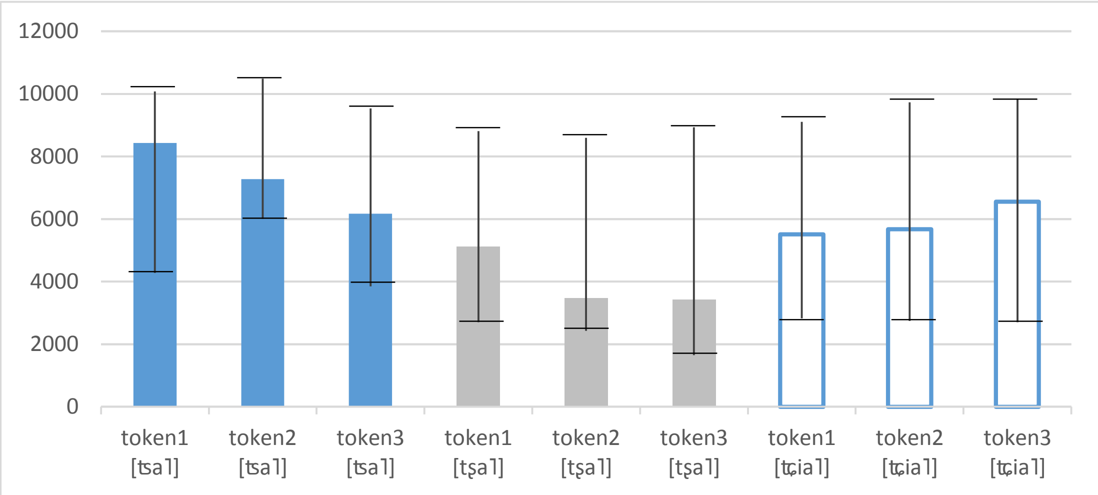


Figure 5. Frequencies in (Hz) of the noise peak (represented by a vertical bar) and noise range (represented by an 'I' line) for the Mandarin unaspirated affricates [ʦ, tʂ, ʨ] in the vowel context [a/ia] produced by Cantonese Male 1.

Figure 5 displays the frequencies of the noise peak and the noise range for three tokens of each of the Mandarin unaspirated affricates [ʦ, tʂ, ʨ] in the vowel context [a/ia] for Cantonese Male 1. Of the three sibilants, the noise peak is obviously lower for the retroflex [tʂ] (3,417-3,472 Hz) than the denti-alveolar [ʦ] (6,174-8,434 Hz) and the alveolo-palatal [ʨ] (5,512-6,559Hz), while it is much higher for one token of [tʂ] (5,126 Hz). The difference of the boundaries in the noise peak is not clear between [ʦ] and [ʨ], and [tʂ] and [ʨ] in some tokens. Regarding minimum value of the noise range, a difference should be taken note of among the three sibilants, which the lowest is also [tʂ] (1,654-2,701 Hz). The minimum value clings to be higher for [ʦ] (3,845-6,011 Hz) than [ʨ] (2,701-2,823 Hz). Compared with the order for the native speaker (Figure 3a), it may be said that Cantonese Male 1 can produce the retroflex [tʂ], but cannot distinguish the denti-alveolar [ʦ] and alveolo-palatal [ʨ], and the retroflex [tʂ] and alveolo-palatal [ʨ].

For Cantonese Male 1, 7,294 Hz and 4,713 Hz are the mean frequencies of three tokens of the noise peak and the minimum value of the noise range for [ʦ], similar to the noise peak of the native in [ʦ] (7,882 Hz), but apparently lower than the minimum value of the native's [ʦ] (6,387 Hz). It is the same case for [ʨ] that the noise peak (5,916 Hz) is similar to that of the native in [ʨ] (6,277 Hz), but the minimum value (2,758 Hz) is distinctly lower than that of the native's [ʨ] (4,884 Hz). Hence, it may suggest that Cantonese Male 1 merges the sibilants [ʦ] into [ʨ], and [ʨ] into [tʂ]. In other words, both [ʦ] and [ʨ] are difficult for Cantonese Male 1.

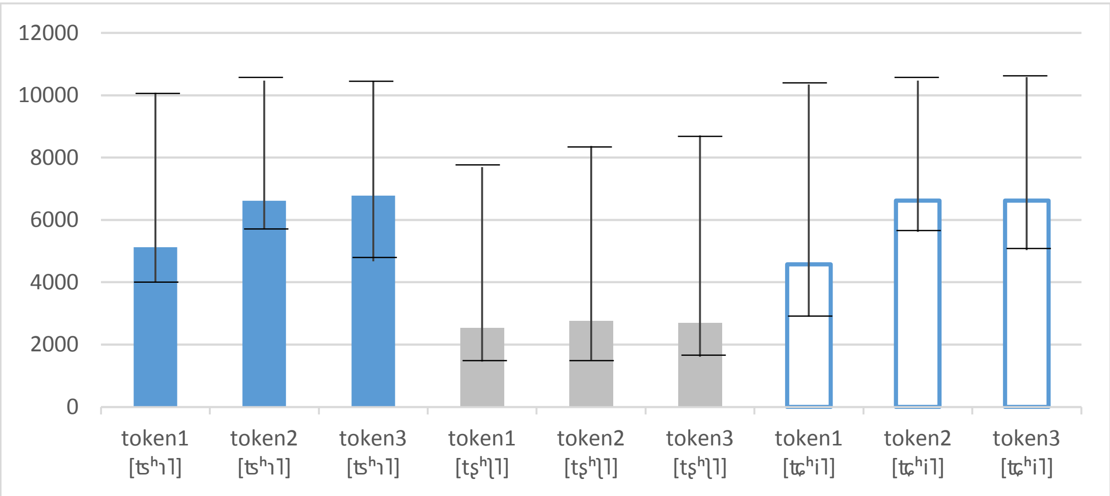


Figure 6. Frequencies in (Hz) of the noise peak (represented by a vertical bar) and noise range (represented by an 'I' line) for the Mandarin aspirated affricates [ʦʰ, tʂʰ, ʨʰ] in the vowel context [i/ɿ/ʅ] produced by Cantonese Male 1.

Figure 6 describes the frequencies of the noise peak and the noise range for three tokens of each of the Mandarin aspirated affricates [ʦʰ, tʂʰ, ʨʰ] in the vowel context [i/ɿ/ʅ] for Cantonese Male 1.Within the three sibilants, the noise peak is clearly lower for the retroflex [tʂʰ] (2,535-2,756 Hz), than the denti-alveolar [ʦʰ] (5,126-6,780 Hz) and the alveolo-palatal [ʨʰ] (4,575-6,615 Hz). The difference in the noise peak between [ʦʰ] and [ʨʰ] is not pronounced. Concerning the noise range, there is a distinct difference among the three sets of sibilants in the minimum value of the noise range, which is also the lowest for [tʂʰ] (1,460-1,606 Hz). The minimum value is prone to be higher for [ʨʰ] (5,037-5,622 Hz) than [ʦʰ] (4,015-5,719 Hz), while it is the lowest for one token of [ʨʰ] (2,896 Hz). By comparison of the noise peak and noise range in a descending order of [ʦʰ] > [ʨʰ] > [tʂʰ] for the native speaker (Figure 3b), it may be considered that Cantonese Male 1 can produce the retroflex [tʂʰ], but cannot differentiate the denti-alveolar [ʦʰ] and alveolo-palatal [ʨʰ].

For Cantonese Male 1, the mean frequencies of three tokens of the noise peak and the minimum value of the noise range for [ʦʰ] are 6,173 Hz and 4,802 Hz, similar to the noise peak of the native in [ʦʰ] (7,882 Hz), but clearly lower than the minimum value of the native's [ʦʰ] (6,387 Hz). 5,935 Hz and 4,518 Hz for [ʨʰ] are similar to those of the native's [ʨʰ] (6,277 Hz and 4,884 Hz). Therefore, it can be said that Cantonese Male 1 merges the sibilant [ʦʰ] into [ʨʰ], which shows that [ʦʰ] is difficult for Cantonese Male 1.

Figure 7 shows the frequencies of the noise peak and the noise range for three tokens of each of the Mandarin aspirated affricates [ʦʰ, tʂʰ, ʨʰ] in the vowel context [a/ia] for Cantonese Male 1. Among the three sibilants, the noise peak is relatively apparently lower for the retroflex [tʂʰ] (2,756-2,866 Hz) than the alveolo-palatal [ʨʰ] (4,244-4,410 Hz) and the denti-alveolar [ʦʰ] (3,031-3,969 Hz). The difference in the noise peak of the three sibilants is not large enough though they are not in overlapping. As for the noise range, a relatively noticeable difference among the three sibilants is in the minimum value of the noise range, which is the highest for [ʨʰ] (2,677-3,748 Hz).

The minimum value of the noise range tends to be lower for [tʂʰ] (1,581-1,752 Hz) than [ʦʰ] (1,630-2,531 Hz). Compared with the order for the native speaker (Figure 3b), it may be said that Cantonese Male 1 can differentiate the retroflex [tʂʰ] but not the denti-alveolar [ʦʰ] and alveolo-palatal [ʨʰ]. Compared with the order for the native speaker (Figure 3b), it may be said that Cantonese Male 1 can differentiate the retroflex [tʂʰ] but not the denti-alveolar [ʦʰ] and alveolo-palatal [ʨʰ].

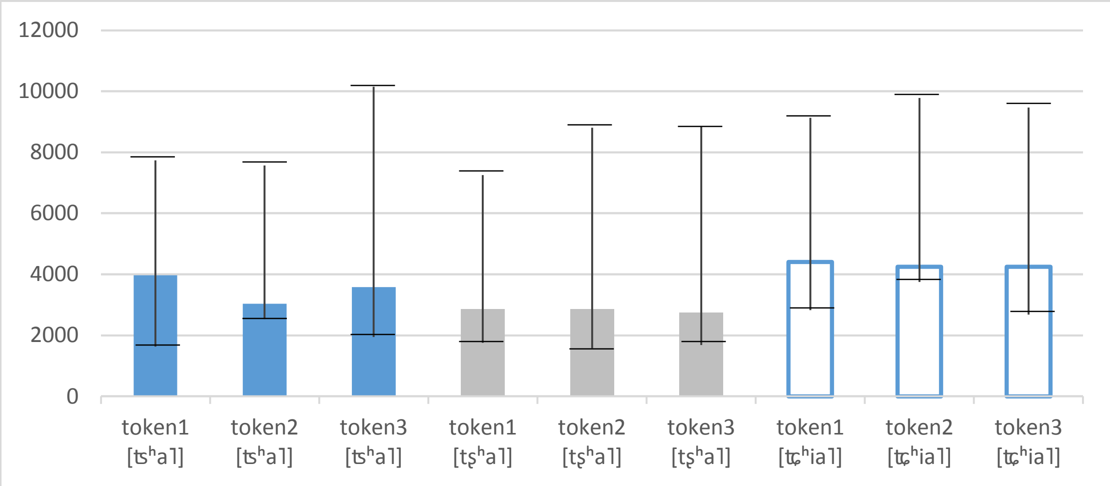


Figure 7. Frequencies in (Hz) of the noise peak (represented by a vertical bar) and noise range (represented by an 'I' line) for the Mandarin aspirated affricates [ʦʰ, tʂʰ, ʨʰ] in the vowel context [a/ia] produced by Cantonese Male 1.

For Cantonese Male 1, the mean frequencies of three tokens of the noise peak and the minimum value of the noise range are 3,527 Hz and 2,036 Hz for [ʦʰ], obviously lower than those of the native's [ʦʰ] (7,882 Hz and 6,387 Hz), but similar to those of the native's [tʂʰ] (3,022 Hz and 1,711 Hz). The noise peak and minimum value of [tʂʰ] are 2,829 Hz and 1,670 Hz, which resembles those of the native in [tʂʰ]. Those of [ʨʰ] are 4,299 Hz and 3,082 Hz, which tends to be closer to those of the native in [tʂʰ] as well. Therefore, it may be suggest that Cantonese Male 1 here merges the sibilants [ʦʰ] and [ʨʰ] into [tʂʰ], treating [tʂʰ] as the free variant for other two sets of sibilants.

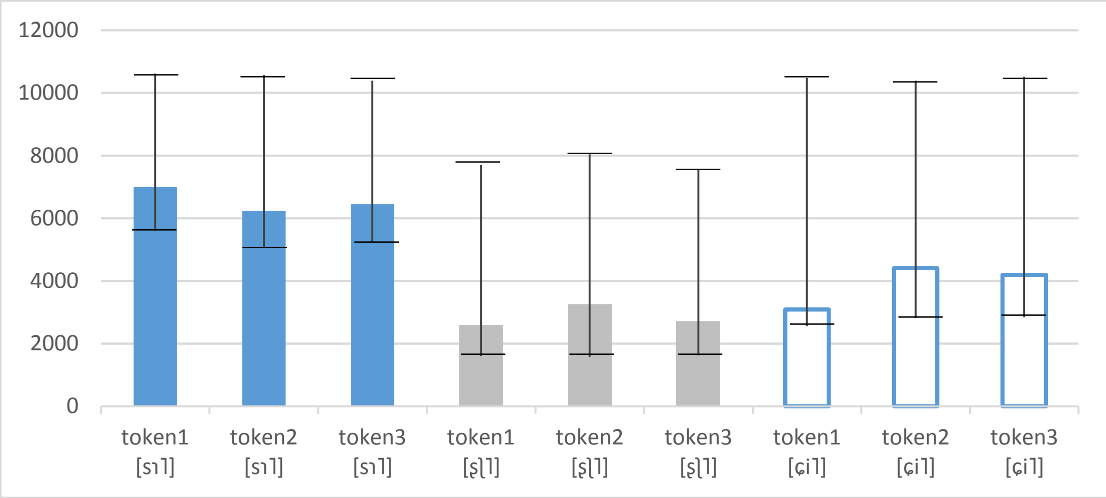


Figure 8. Frequencies in (Hz) of the noise peak (represented by a vertical bar) and noise range (represented by an 'I' line) for the Mandarin fricatives [s, ʂ, ɕ] in the vowel context [i,ɿ,ʅ] produced by Cantonese Male 1.

Figure 8 shows the frequencies of the noise peak and the noise range for three tokens of each of the Mandarin fricatives [s, ʂ, ɕ] in the vowel context [i,ɿ,ʅ] for Cantonese Male 1. Of the three sibilants, the noise peak is apparently higher for the denti-alveolar [s] (6,229-7,000 Hz) than the alveolo-palatal [ɕ] (3,087-4,410 Hz) and the retroflex [ʂ] (2,590-3,252 Hz). The difference of the boundaries in the noise peak is not clear between [ɕ] and [ʂ]. As for the noise range, a noticeable difference among the three sibilants is in the minimum value of the noise range, which is also the highest for [s] (5,086-5,622 Hz). The minimum value of the noise range is cling to be lower for [ʂ] (1,606-1,654 Hz) than [ɕ] (2,579-2,871 Hz). By comparison of the noise peak and noise range in a descending order of [s] > [ɕ] > [ʂ] for the native speaker (Figure 3c), it may be considered that Cantonese Male 1 can distinguish the retroflex [ʂ], the denti-alveolar [s] and alveolo-palatal [ɕ].

For Cantonese Male 1, the mean frequencies of three tokens of the noise peak and the minimum value of the noise range are 6,559 Hz and 5,329 Hz for [s], which are similar to those of the native's [ɕ] (6,277 Hz and 4,884 Hz) but not [s] (7,882 Hz and 6,387 Hz). The noise peak and the minimum value for [ɕ] are 3,895 Hz and 2,765 Hz, which resemble those of the native's [ʂ] (3,022 Hz and 1,711 Hz) but not [ɕ] (6,277 Hz and 4,884 Hz). Thus, it may suggest that the Cantonese Male 1 merges the sibilants [s] into [ɕ], and [ɕ] into [ʂ].

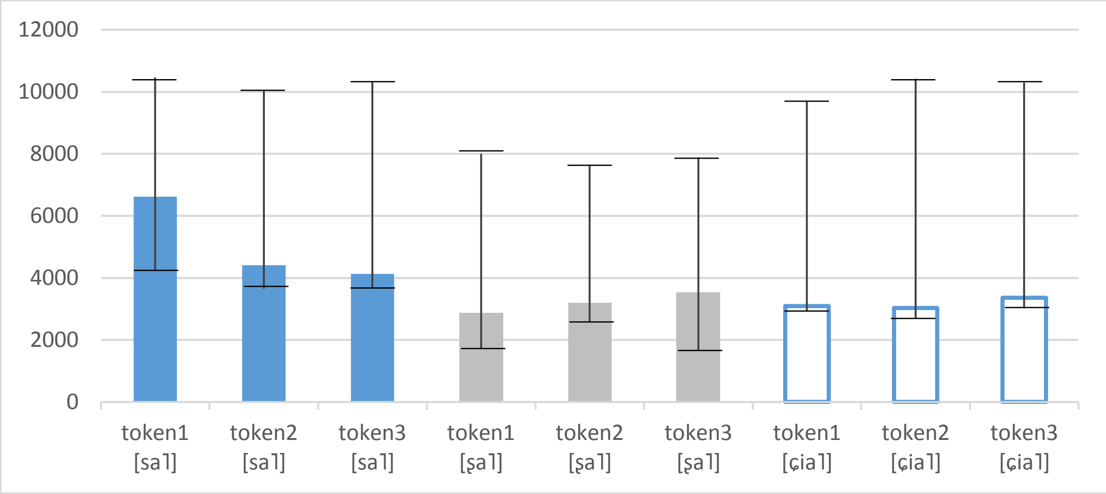


Figure 9. Frequencies in (Hz) of the noise peak (represented by a vertical bar) and noise range (represented by an 'I' line) for the Mandarin fricatives [s, ʂ, ɕ] in the vowel context [a/ia] produced by Cantonese Male 1.

Figure 9 reveals the frequencies of the noise peak and the noise range for three tokens of each of the Mandarin fricatives [s, ʂ, ɕ] in the vowel context [a/ia] for Cantonese Male 1. Within the three sibilants, the noise peak is apparently the highest for one token in the denti-alveolar [s] (6,615 Hz). The noise peaks are 4,134-4,410 Hz for [s], 3,031-3,362 Hz for [ɕ], and 2,866-3,528 Hz for [ʂ], which are very similar that the differences between [s], [ɕ] and [ʂ] are not obvious. Concerning the noise range, there is a difference among the three sets of sibilants in the minimum value of the noise range, which is also the lowest for [ʂ] (1,679-2,604 Hz). The minimum value is prone to be higher for [s] (3,650-4,259 Hz) than [ɕ] (2,944-3,017 Hz). However, their differences are not large enough. Compared with the order for the native speaker (Figure 3c), it may be considered that Cantonese Male 1 can produce the retroflex [ʂ], but cannot differentiate the denti-alveolar [s] and alveolo-palatal [ɕ].

For Cantonese Male 1, the mean frequencies of three tokens of the noise peak and the minimum value of the noise range for [s] are 5,053 Hz and 3,869 Hz, which are apparently lower than those of the native's [s] (7,882 Hz and 6,387 Hz), and similar to native's [ɕ] (6,277 Hz and 4,884 Hz). The noise peak and the minimum value for [ɕ] are 3,160 Hz and 2,887 Hz, which are much lower than those of the native's [ɕ] but similar to native's [ʂ] (3,022 Hz and 1,711 Hz). Thus, it may be considered that Cantonese Male 1 merges [s] into [ɕ], and [ɕ] into [ʂ].

3.3 Cantonese Male 2

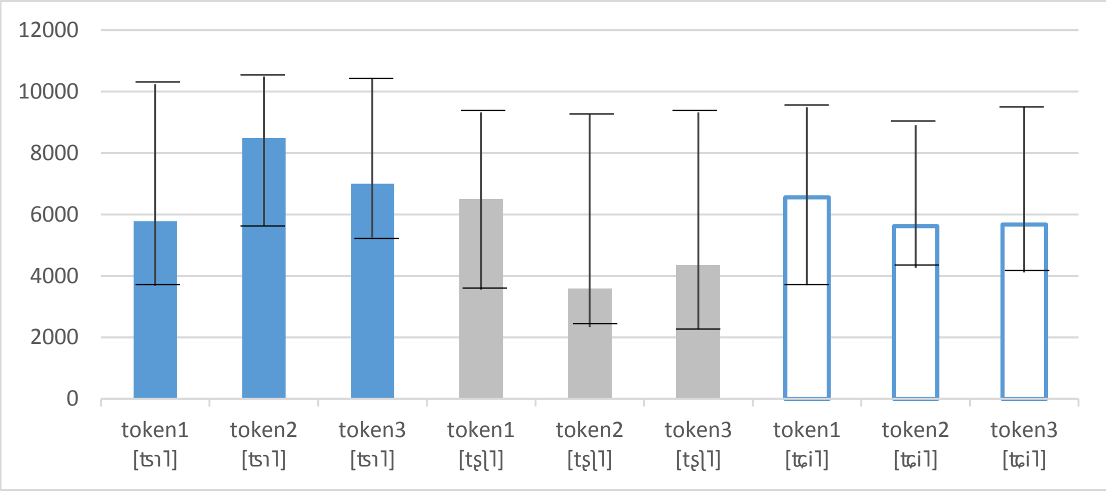


Figure 10. Frequencies in (Hz) of the noise peak (represented by a vertical bar) and noise range (represented by an 'I' line) for the Mandarin unaspirated affricates [ʦ, tʂ, ʨ] in the vowel context [i/ɿ/ʅ] produced by Cantonese Male 2.

Figure 10 shows the frequencies (in Hz) of the noise peak and the noise range for three tokens of each of the Mandarin unaspirated affricates [ʦ, tʂ, ʨ] in the vowel context [i/ɿ/ʅ] for Cantonese Male 2. Among the three sibilants, the noise peak is noticeably lower for the retroflex [tʂ] (3,583-4,354 Hz) than the denti-alveolar [ʦ] (7,000-8,489 Hz) and the alveolo-palatal [ʨ] (5,622-6,559 Hz), while it is higher for one token of [tʂ] (6,504 Hz) and lower for one token of [ʦ] (5,788 Hz). Between [ʦ] and [ʨ], the difference in the noise peak is not pronounced. As for the noise range, a noticeable difference among the three sibilants is in the minimum value of the noise range, which is also the lowest for [tʂ] (2,263-3,553 Hz). The minimum value of the noise range tends to be higher for [ʦ] (3,675-5,622 Hz) than [ʨ] (3,699-4,259 Hz). By comparison of the noise peak and noise range in a descending order of [ʦ] > [ʨ] > [tʂ] for the native speaker (Figure 3a), it may be considered that Cantonese Male 2 can differentiate most of the sibilants here, except some tokens.

For Cantonese Male 2, the mean frequencies of three tokens of the noise peak and the minimum value of the noise range are 7,092 Hz and 4,851 Hz for [ʦ], similar to the noise peak of the native's [ʦ] (7,882 Hz), and apparently lower than minimum value of the native's [ʦ] (6,387 Hz). The noise peak and the minimum value for [tʂ] are 4,813 Hz and 2,717 Hz, obviously higher than the peak of the native's [tʂ] (3,022 Hz) and similar to the minimum value of the native's [tʂ] (1,711 Hz). Those for [ʨ] are 5,952 Hz and 4,023 Hz, similar to those of the native's [ʨ] (6,277 Hz and 4,884 Hz). Thus, it may suggest that Cantonese Male 2 merges the sibilants [ʦ] and [tʂ] into [ʨ].

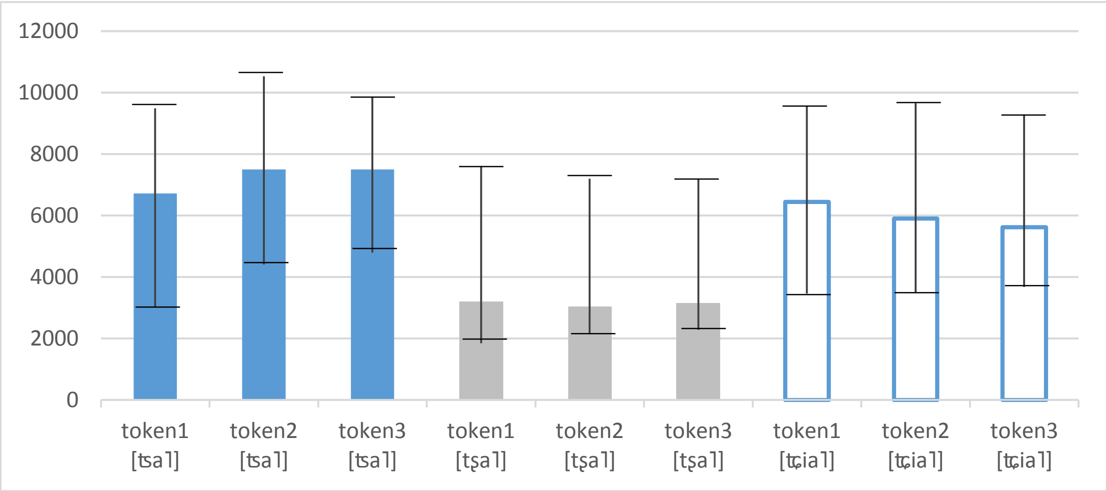


Figure 11. Frequencies in (Hz) of the noise peak (represented by a vertical bar) and noise range (represented by an 'I' line) for the Mandarin unaspirated affricates [ʦ, tʂ, ʨ] in the vowel context [a/ia] produced by Cantonese Male 2.

Figure 11 displays the frequencies of the noise peak and the noise range for three tokens of each of the Mandarin unaspirated affricates [ʦ, tʂ, ʨ] in the vowel context [a/ia] for Cantonese Male 2. Of the three sibilants, the noise peak is obviously lower for the retroflex [tʂ] (3,031-3,142 Hz) than the denti-alveolar [ʦ] (6,725-7,497 Hz) and the alveolo-palatal [ʨ] (5,622-6,449 Hz).The difference in the noise peak is apparent for the sibilants. Regarding minimum value of the noise range, a difference should be paid attention to among the three sibilants, which the lowest is also [tʂ] (1,849-2,287 Hz). The minimum value clings to be higher for [ʦ] (4,405-4,794 Hz) than [ʨ] (3,455-3,675 Hz), while it is much lower for one token of [ʦ] (3,017 Hz). Compared with the order for the native speaker (Figure 3a), it may be said that Cantonese Male 2 can differentiate the retroflex [tʂ], denti-alveolar [ʦ] and alveolo-palatal [ʨ], except one token in [ʦ].

For Cantonese Male 2, 7,239 Hz and 4,072 Hz are the mean frequencies of three tokens of the noise peak and the minimum value of the noise range for [ʦ], similar to the noise peak of the native in [ʦ] (7,882 Hz), but apparently lower than the minimum value of the native's [ʦ] (6,387 Hz). The noise peak and the minimum value for [ʨ] are 5,989 Hz and 3,544 Hz, similar to those of the native's [ʨ] (6,277 Hz and 4,884 Hz). Hence, it may be considered that Cantonese Male 2 merges [ʦ] into [ʨ].

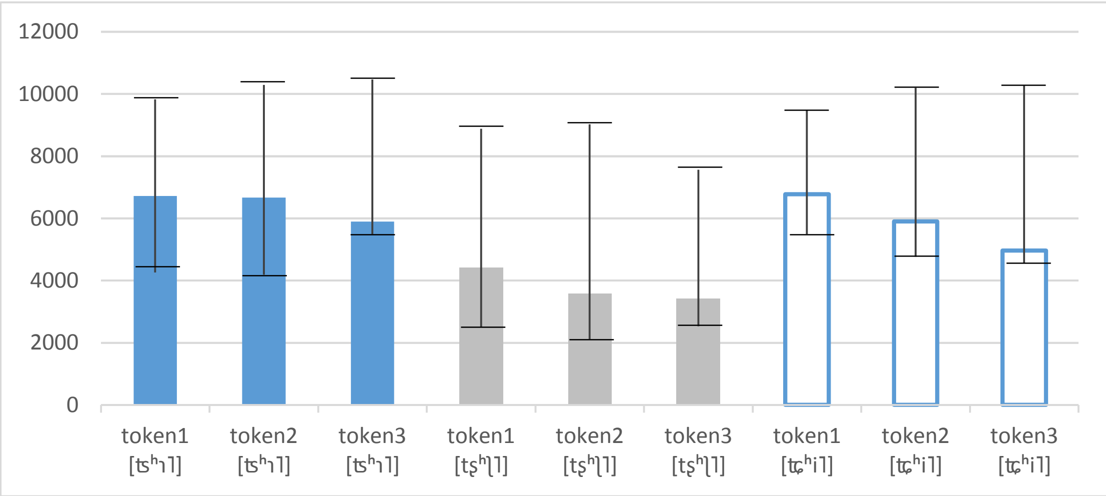


Figure 12. Frequencies in (Hz) of the noise peak (represented by a vertical bar) and noise range (represented by an 'I' line) for the Mandarin aspirated affricates [ʦʰ, tʂʰ, ʨʰ] in the vowel context [i/ɿ/ʅ] produced by Cantonese Male 2.

Figure 12 describes the frequencies of the noise peak and the noise range for three tokens of each of the Mandarin aspirated affricates [ʦʰ, tʂʰ, ʨʰ] in the vowel context [i/ɿ/ʅ] for Cantonese Male 2.Within the three sibilants, the noise peak is clearly lower for the retroflex [tʂʰ] (3,417-4,410 Hz), than the denti-alveolar [ʦʰ] (5,898-6,725 Hz) and the alveolo-palatal [ʨʰ] (5,898-6,780 Hz), while it is much lower in one token of [ʨʰ] (4,961 Hz). The difference in the noise peak between [ʦʰ] and [ʨʰ] is not pronounced. Concerning the noise range, there is a distinct difference among the three sets of sibilants in the minimum value of the noise range, which is also the lowest for [tʂʰ] (2,117-2,531 Hz). The minimum value is similar for [tɕʰ] (4,551-5,500 Hz) and [tsʰ] (4,186-5,500 Hz). By comparison of the noise peak and noise range in a descending order of [tsʰ] > [tɕʰ] > [tʂʰ] for the native speaker (Figure 3b), it may be considered that Cantonese Male 2 can produce the retroflex [tʂʰ], but cannot differentiate the denti-alveolar [tsʰ] and alveolo-palatal [tɕʰ].

For Cantonese Male 2, the mean frequencies of three tokens of the noise peak and the minimum value of the noise range for [tsʰ] are 6,431 Hz and 4,648 Hz, much lower than those of the native's [tsʰ] (7,882 Hz and 6,387 Hz), and 5,879 Hz and 4,940 Hz for [tɕʰ] are similar to those of the native's [tɕʰ] (6,277 Hz and 4,884 Hz). Therefore, it can be said that Cantonese Male 2 merges the sibilant [tsʰ] into [tɕʰ], which shows that [tsʰ] is difficult for Cantonese Male 2.

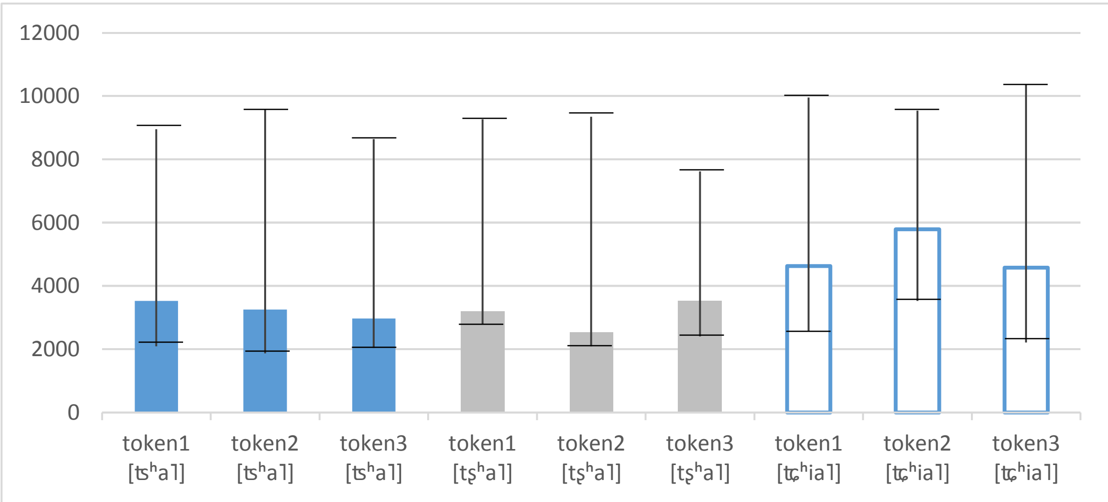


Figure 13. Frequencies in (Hz) of the noise peak (represented by a vertical bar) and noise range (represented by an 'I' line) for the Mandarin aspirated affricates [tsʰ, tʂʰ, tɕʰ] in the vowel context [a/ia] produced by Cantonese Male 2.

Figure 13 shows the frequencies of the noise peak and the noise range for three tokens of each of the Mandarin aspirated affricates [tsʰ, tʂʰ, tɕʰ] in the vowel context [a/ia] for Cantonese Male 2. Among the three sibilants, the noise peak is relatively apparently higher for the alveolo-palatal [ʨʰ] (4,575-5,788 Hz) than the retroflex [tʂʰ] (2,535-3,528 Hz) and the denti-alveolar [ʦʰ] (2,976-3,528 Hz). The difference between [tʂʰ] and [ʦʰ] is clear. As for the noise range, a relatively noticeable difference among the three sibilants is in the minimum value of the noise range, which is the lowest for [ʦʰ] (1,874-2,093 Hz). The minimum value of the noise range tends to be higher for [ʨʰ] (2,214-3,528 Hz) than [tʂʰ] (2,093-2,798 Hz). Compared with the order for the native speaker (Figure 3b), it may be said that Cantonese Male 2 cannot differentiate the retroflex [tʂʰ], the denti-alveolar [ʦʰ] and alveolo-palatal [ʨʰ].

For Cantonese Male 2, the mean frequencies of three tokens of the noise peak and the minimum value of the noise range are 3,252Hz and 2,003 Hz for [ʦʰ], obviously lower than those of the native's [ʦʰ] (7,882 Hz and 6,387 Hz), but similar to those of the native's [tʂʰ] (3,022 Hz and 1,711 Hz). The noise peak and minimum value of [tʂʰ] are 3,086 Hz and 2,433 Hz, which resembles those of the native in [tʂʰ]. Those of [ʨʰ] are 4,997 Hz and 2,765 Hz, apparently lower than those of the native's [ʨʰ] (6,277 Hz and 4,884 Hz). Therefore, it may be suggest that Cantonese Male 2 here merges the sibilants [ʦʰ] and [ʨʰ] into [tʂʰ], treating [tʂʰ] as the free variant for other two sets of sibilants.

Figure 14 shows the frequencies of the noise peak and the noise range for three tokens of each of the Mandarin fricatives [s, ʂ, ɕ] in the vowel context [i,ɿ,ʅ] for Cantonese Male 2. Of the three sibilants, the noise peak is apparently higher for the denti-alveolar [s] (7,386-8,544 Hz) than the alveolo-palatal [ɕ] (6,615-6,835 Hz) and the retroflex [ʂ] (3,417-4,630 Hz), while it is much higher for one token of [ʂ] (5,843 Hz). The difference of the boundaries in the noise peak is clear for the three sibilants, except one token of [ʂ]. As for the noise range, a noticeable difference among the three sibilants is in the minimum value of the noise range, which is also the highest for [s]

(5,086-5,622 6,035 Hz), while two tokens of [s] are much lower (4,502-4,624 Hz), similar to that of [ɕ] (4,405-4,648 Hz). The minimum value of the noise range of [ʂ] is the lowest (2,677-2,823 Hz) with a token higher (4,478 Hz). By comparison of the noise peak and noise range in a descending order of [s] > [ɕ] > [ʂ] for the native speaker (Figure 3c), it may be considered that Cantonese Male 2 can produce alveolo-palatal [ɕ], but cannot distinguish the retroflex [ʂ] and the denti-alveolar [s].

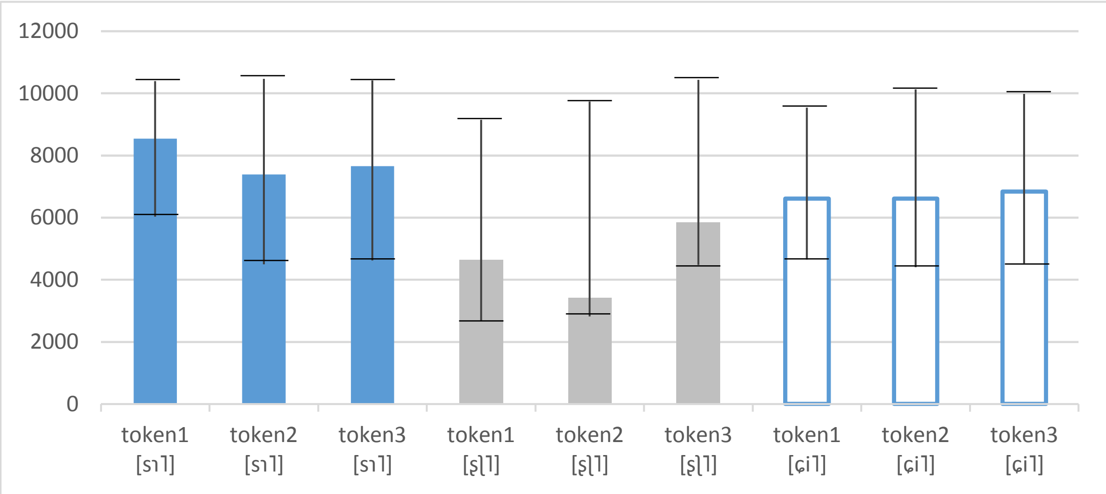


Figure 14. Frequencies in (Hz) of the noise peak (represented by a vertical bar) and noise range (represented by an ‘I’ line) for the Mandarin fricatives [s, ʂ, ɕ] in the vowel context [i,ɿ,ʅ] produced by Cantonese Male 2.

For Cantonese Male 2, the mean frequencies of three tokens of the noise peak and the minimum value of the noise range are 7,864 Hz and 5,053 Hz for [s], similar to the noise peak of the native’s [s] (7,882 Hz) but lower than the minimum value of the native (6,387 Hz). The noise peak and the minimum value for [ʂ] are 4,630 Hz and 3,326 Hz, which are higher those of the native’s [ʂ] (3,022 Hz and 1,711 Hz). Thus, it may suggest that the Cantonese Male 2 merges the sibilants [s] and [ʂ] into [ɕ].

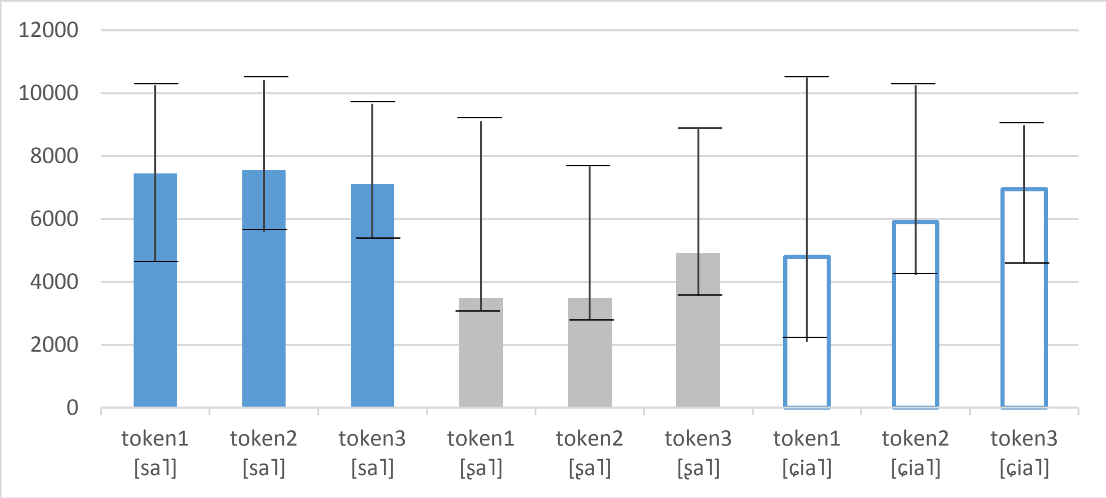


Figure 15. Frequencies in (Hz) of the noise peak (represented by a vertical bar) and noise range (represented by an 'I' line) for the Mandarin fricatives [s, ʂ, ɕ] in the vowel context [a/ia] produced by Cantonese Male 2.

Figure 15 reveals the frequencies of the noise peak and the noise range for three tokens of each of the Mandarin fricatives [s, ʂ, ɕ] in the vowel context [a/ia] for Cantonese Male 2. Within the three sibilants, the noise peak is apparently lower for the retroflex [ʂ] (3,472-3,472 Hz) than the denti-alveolar [s] (7,111-7,552 Hz) and the alveolo-palatal [ɕ] (5,898-6,945 Hz), while it is much higher for one token of [ʂ] (4,906 Hz) and lower for one token of [ɕ] (4,795 Hz). The difference is obvious for the three sibilants, except the two mentioned tokens. Concerning the noise range, there is a difference among the three sets of sibilants in the minimum value of the noise range, which is also the lowest for [ʂ] (2,798-3,553 Hz). The minimum value is prone to be higher for [s] (4,672-5,573 Hz) than [ɕ] (4,210-4,575 Hz), while it is much lower for one token of [ɕ] (2,093 Hz). However, their differences are not large enough. Compared with the order for the native speaker (Figure 3c), it may be considered that Cantonese Male 2 can differentiate the retroflex [ʂ], the denti-alveolar [s] and alveolo-palatal [ɕ], except a few tokens.

For Cantonese Male 2, the mean frequencies of three tokens of the noise peak and the minimum value of the noise range for [s] are 7,368 Hz and 5,207 Hz, which are just lower than the minimum value of the native's [s] (6,387 Hz) but similar to the noise peak of the native (7,882 Hz). The noise peak and the minimum value for [ɕ] are 5,879 Hz and 3,626 Hz, while those for [ʂ] are 3,950 Hz and 3,147 Hz. Both [ɕ] and [ʂ] are similar to their native's [ɕ] (6,277 Hz and 4,884 Hz) and [ʂ] (3,022 Hz and 1,711 Hz). Thus, it may be considered that Cantonese Male 2 merges [s] into [ɕ].

3.4 Cantonese Female 1

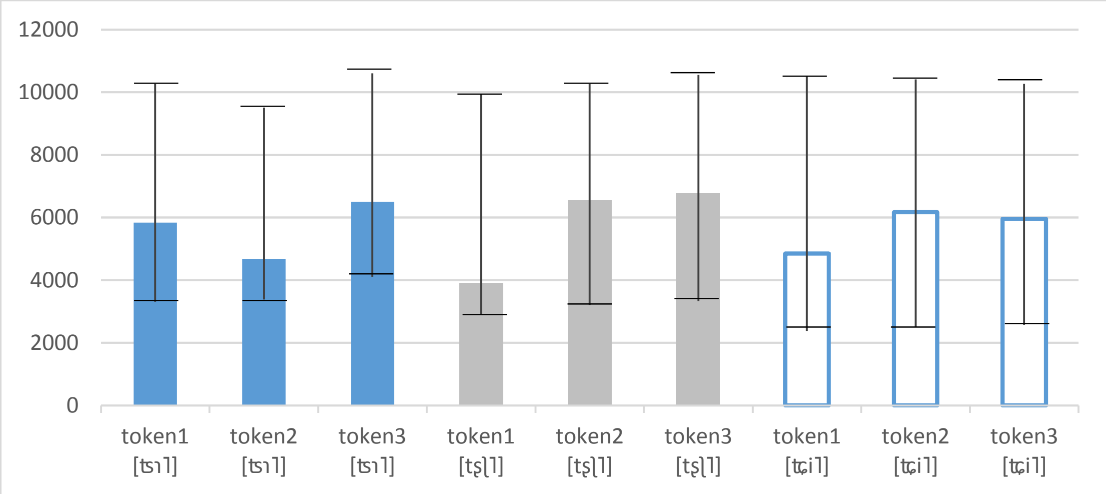


Figure 16. Frequencies in (Hz) of the noise peak (represented by a vertical bar) and noise range (represented by an 'I' line) for the Mandarin unaspirated affricates [ts, tʂ, tɕ] in the vowel context [i/ɿ/ʅ] produced by Cantonese Female 1.

Figure 16 shows the frequencies (in Hz) of the noise peak and the noise range for three tokens of each of the Mandarin unaspirated affricates [ts, tʂ, tɕ] in the vowel context [i/ɿ/ʅ] for Cantonese Female 1. Among the three sibilants, the noise peak is noticeably higher for the retroflex [tʂ] (6,559-6,780 Hz) than the denti-alveolar [ts] (4,685-6,504 Hz) and the alveolo-palatal [tɕ] (4,851-6,174 Hz), while it is much lower for one token of [tʂ] (3,913 Hz). Between [ts] and [tɕ], the difference in the noise peak is not pronounced. As for the noise range, a noticeable difference among the three sibilants is in the minimum value of the noise range, which is also the lowest for [tɕ] (2,385-2,579 Hz). The minimum value of the noise range tends to be higher for [ts] (3,309-4,113 Hz) than [tʂ] (2,920-3,334 Hz). By comparison of the noise peak and noise range in a descending order of [ts] > [tɕ] > [tʂ] for the native speaker (Figure 3a), it may be considered that Cantonese Female 1 cannot distinguish the retroflex [tʂ], the denti-alveolar [ts] and alveolo-palatal [tɕ].

For Cantonese Female 1, the mean frequencies of three tokens of the noise peak and the minimum value of the noise range are 5,677 Hz and 3,601 Hz for [ʦ], apparently lower than those of the native's [ʦ] (7,882Hz and 6,387 Hz), and 5,659 Hz and 2,490 Hz for [ʨ], also lower than those of the native's [ʨ] (6,277 Hz and 4,884 Hz). The noise peak and minimum value for [tʂ] are 5,750 Hz and 3,155 Hz, which are obviously higher than those of the native's [tʂ] (3,022 Hz and 1,711 Hz). Thus, it may suggest that Cantonese Female 1 merges the sibilants [ʦ] and [tʂ] into [ʨ], and [ʨ] into [tʂ].

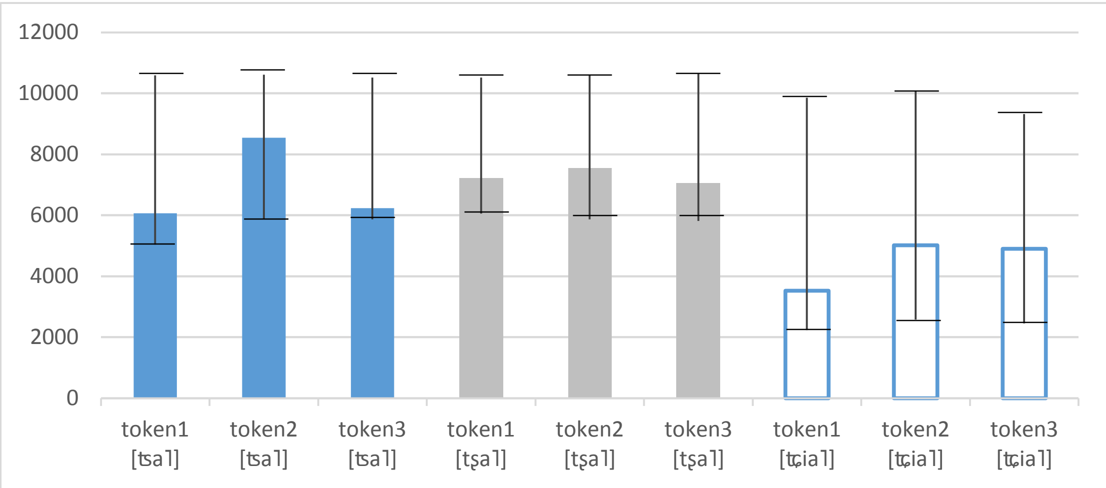


Figure 17. Frequencies in (Hz) of the noise peak (represented by a vertical bar) and noise range (represented by an 'I' line) for the Mandarin unaspirated affricates [ʦ, tʂ, ʨ] in the vowel context [a/ia] produced by Cantonese Female 1.

Figure 17 displays the frequencies of the noise peak and the noise range for three tokens of each of the Mandarin unaspirated affricates [ʦ, tʂ, ʨ] in the vowel context [a/ia] for Cantonese Female 1. Of the three sibilants, the noise peak is obviously lower for the alveolo-palatal [ʨ] (3,528-5,016 Hz) than the retroflex [tʂ] (7,056-7,552 Hz) and the denti-alveolar [ʦ] (6,063-8,544 Hz). The difference in the noise peak is not clear between [ʦ] and [tʂ]. Regarding minimum value of the noise range, a difference should be taken note of among the three sibilants, which the lowest is also [ʨ] (2,239-

2,579 Hz). The minimum value of [ʦ] (5,037-5,889 Hz) is similar to that of [ʈʂ] (5,816-6,060 Hz). Compared with the order for the native speaker (Figure 3a), it may be said that Cantonese Female 1 cannot distinguish the denti-alveolar [ʦ], alveolo-palatal [ʨ], and the retroflex [ʈʂ].

For Cantonese Female 1, 6,945 Hz and 5,597 Hz are the mean frequencies of three tokens of the noise peak and the minimum value of the noise range for [ʦ], similar to those of the native in [ʦ] (7,882 Hz and 6,387 Hz). The noise peak and the minimum value for [ʈʂ] are 7,276 Hz and 5,913 Hz, which are apparently higher than those of the native in [ʈʂ] (3,022 Hz and 1,711 Hz). Those for [ʨ] are 4,483 Hz and 2,425 Hz, which are distinctively lower than those of the native in [ʨ] (6,277 Hz and 4,884 Hz). Hence, it may suggest that Cantonese Female 1 merges the sibilants [ʈʂ] into [ʦ], and [ʨ] into [ʈʂ].

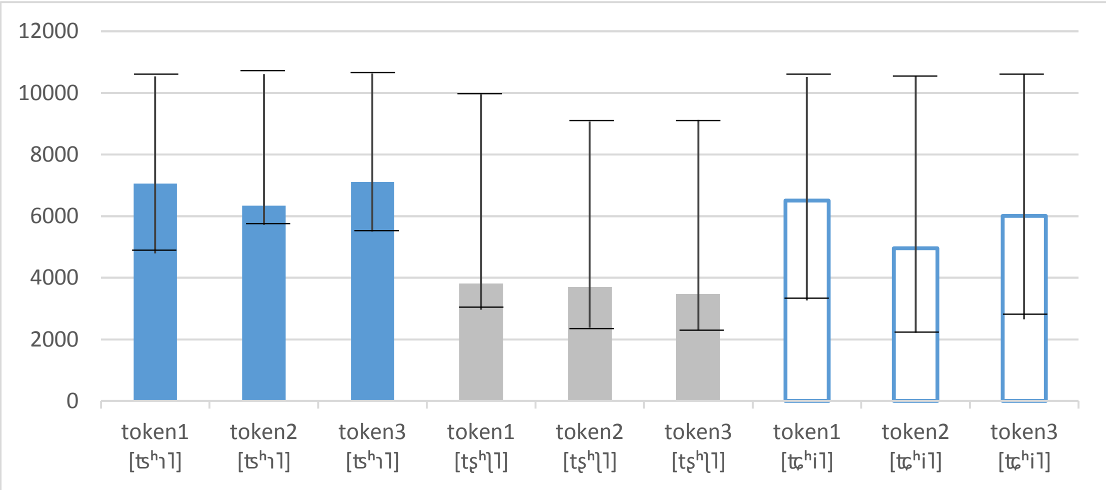


Figure 18. Frequencies in (Hz) of the noise peak (represented by a vertical bar) and noise range (represented by an 'I' line) for the Mandarin aspirated affricates [ʦʰ, ʈʂʰ, ʨʰ] in the vowel context [i/ɿ/ʅ] produced by Cantonese Female 1.

Figure 18 describes the frequencies of the noise peak and the noise range for three tokens of each of the Mandarin aspirated affricates [ʦʰ, ʈʂʰ, ʨʰ] in the vowel context [i/ɿ/ʅ] for Cantonese Female 1. Within the three sibilants, the noise peak is clearly lower for the retroflex [tʂʰ] (3,472-3,803 Hz), than the denti-alveolar [ʦʰ] (6,339-7,111 Hz) and the alveolo-palatal [ʨʰ] (6,008-6,504Hz), while it is much lower for one token in [ʨʰ] (4,961 Hz). The difference in the noise peak between [ʦʰ] and [ʨʰ] is not pronounced. Concerning the noise range, there is a distinct difference among the three sets of sibilants in the minimum value of the noise range, which the highest is [ʦʰ] (4,794-5,719 Hz). The minimum value is prone to be lower for [ʨʰ] (2,214-3,261 Hz) and [tʂʰ] (2,312-2,969 Hz). By comparison of the noise peak and noise range in a descending order of [ʦʰ] > [ʨʰ] > [tʂʰ] for the native speaker (Figure 3b), it may be considered that Cantonese Female 1 cannot differentiate the denti-alveolar [ʦʰ], alveolo-palatal [ʨʰ] and the denti-alveolar [ʦʰ].

For Cantonese Female 1, the mean frequencies of three tokens of the noise peak and the minimum value of the noise range for [ʦʰ] are 6,835 Hz and 5,337 Hz, clearly lower those of the native's [ʦʰ] (7,882 Hz and 6,387 Hz), and 5,824 Hz and 2,709 Hz for [ʨʰ] are similar to the noise peak of the native's [ʨʰ] (6,277 Hz) but lower than the minimum value of the native (4,884 Hz). The noise peak and minimum value for [tʂʰ] are 3,656 Hz and 2,555 Hz, which are similar to those of the native in [tʂʰ] (3,022 Hz and 1,711 Hz). Therefore, it can be said that Cantonese Female 1 merges the sibilant [ʦʰ] into [ʨʰ], and [ʨʰ] into [tʂʰ].

Figure 19 shows the frequencies of the noise peak and the noise range for three tokens of each of the Mandarin aspirated affricates [ʦʰ, tʂʰ, ʨʰ] in the vowel context [a/ia] for Cantonese Female 1. Among the three sibilants, the noise peak is apparently lower for the alveolo-palatal [ʨʰ] (4961-5016 Hz) than the retroflex [tʂʰ] (7,000-7,111 Hz) and the denti-alveolar [ʦʰ] (6,118-8,379 Hz). The difference in the noise peak between [tʂʰ] and [ʦʰ] is not clear. As for the noise range, a relatively noticeable difference among the three sibilants is in the minimum value of the noise range, which is the lowest for [ʨʰ] (2,287-2,823 Hz). The minimum value of the noise range tends to be higher for [tʂʰ] (3,455-4,843 Hz) than [ʦʰ] (4,745-5,573 Hz). Compared with the order for the native speaker (Figure 3b), it may be said that Cantonese Female 1 cannot differentiate the retroflex [tʂʰ], the denti-alveolar [ʦʰ] and alveolo-palatal [ʨʰ].

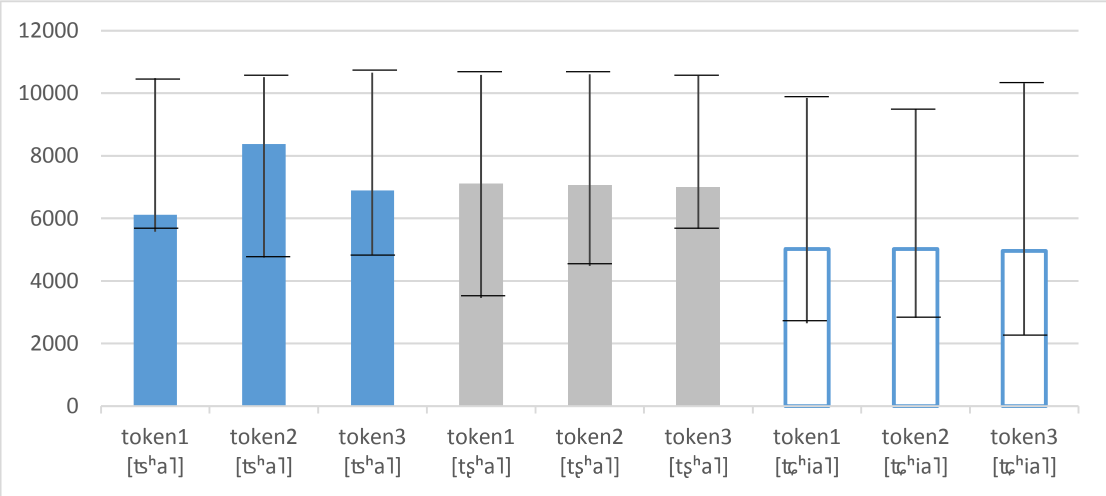


Figure 19. Frequencies in (Hz) of the noise peak (represented by a vertical bar) and noise range (represented by an 'I' line) for the Mandarin aspirated affricates [ʦʰ, tʂʰ, ʨʰ] in the vowel context [a/ia] produced by Cantonese Female 1.

For Cantonese Female 1, the mean frequencies of three tokens of the noise peak and the minimum value of the noise range are 7,129 Hz and 5,053 Hz for [ʦʰ], similar to the noise peak of the native's [ʦʰ] (7,882 Hz), and obviously lower than the minimum value of the native (6,387 Hz). The noise peak and minimum value of [tʂʰ] are 7,055 Hz and 4,534 Hz, which are apparently much higher than those of the native in [tʂʰ] (3,022 Hz and 1,711 Hz). Those of [ʨʰ] are 4,997 Hz and 2,587 Hz, which are much lower than those of the native in [ʨʰ] (6,277 Hz and 4,884 Hz). Therefore, it may be suggest that Cantonese Female 1 here merges the sibilants [ʦʰ] into [ʨʰ], [tʂʰ] into [ʦʰ] and [ʨʰ], and [ʨʰ] into [tʂʰ].

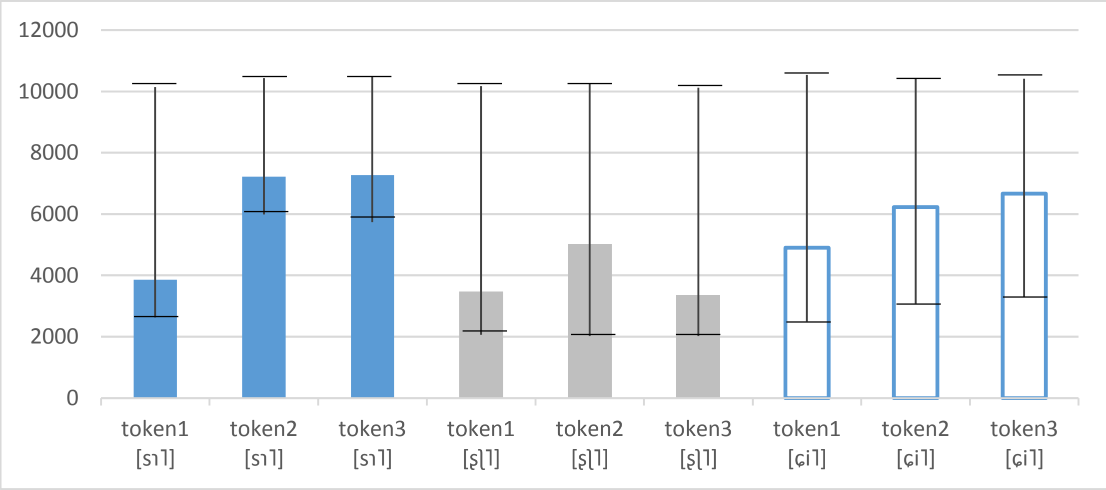


Figure 20. Frequencies in (Hz) of the noise peak (represented by a vertical bar) and noise range (represented by an 'I' line) for the Mandarin fricatives [s, ʂ, ɕ] in the vowel context [i,ɿ,ʅ] produced by Cantonese Female 1.

Figure 20 shows the frequencies of the noise peak and the noise range for three tokens of each of the Mandarin fricatives [s, ʂ, ɕ] in the vowel context [i,ɿ,ʅ] for Cantonese Female 1. Of the three sibilants, the noise peak is apparently higher for the denti-alveolar [s] (7,221-7,276 Hz) than the alveolo-palatal [ɕ] (4,906-6,670 Hz) and the retroflex [ʂ] (3,362-5,016 Hz), while it is much lower for one token of [s] (3,858 Hz). The difference in the noise peak is clear between the three sibilants, except the token of [s]. As for the noise range, a noticeable difference among the three sibilants is in the minimum value of the noise range, which is also the highest for [s] (5,743-5,987 Hz), while it is much lower for one token of [s] (2,628 Hz). The minimum value of the noise range is cling to be lower for [ʂ] (2,020-2,068 Hz) than [ɕ] (2,506-3,285 Hz). By comparison of the noise peak and noise range in a descending order of [s] > [ɕ] > [ʂ] for the native speaker (Figure 3c), it may be considered that Cantonese Female 1 can distinguish the retroflex [ʂ], the denti-alveolar [s] and alveolo-palatal [ɕ].

For Cantonese Female 1, the mean frequencies of three tokens of the noise peak and the minimum value of the noise range are 6,118 Hz and 4,786 Hz for [s], which are similar to those of the native's [ɕ] (6,277 Hz and 4,884 Hz) but not [s] (7,882

Hz and 6,387 Hz). The noise peak and the minimum value for [ɕ] are 5,935 Hz and 2,960 Hz, which are similar to the noise peak of the native's [ɕ] (6,277 Hz) and resemble the minimum value of the native's [ʂ] (1,711 Hz). Thus, it may suggest that the Cantonese Female 1 merges the sibilants [s] into [ɕ], and [ɕ] into [ʂ].

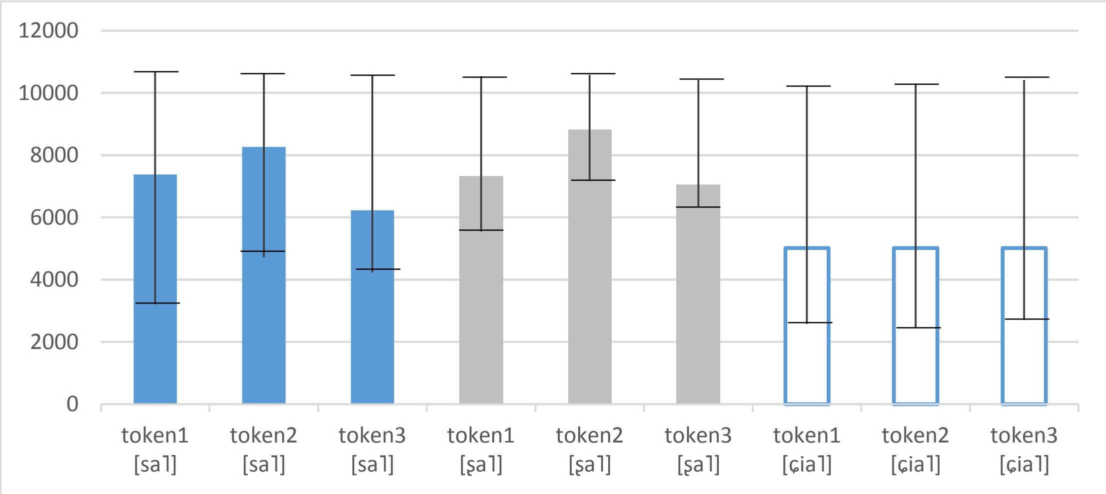


Figure 21. Frequencies in (Hz) of the noise peak (represented by a vertical bar) and noise range (represented by an 'I' line) for the Mandarin fricatives [s, ʂ, ɕ] in the vowel context [a/ia] produced by Cantonese Female 1.

Figure 21 reveals the frequencies of the noise peak and the noise range for three tokens of each of the Mandarin fricatives [s, ʂ, ɕ] in the vowel context [a/ia] for Cantonese Female 1. Within the three sibilants, the noise peak for the alveolo-palatal [ɕ] (5,016-5,016 Hz) is lower than the denti-alveolar [s] (6,229-8,268 Hz) and the retroflex [ʂ] (7,056-8,820 Hz). The difference between [s] and [ʂ] are not obvious. Concerning the noise range, there is a difference among the three sets of sibilants in the minimum value of the noise range, which is also the lowest for [ɕ] (2,458-2,701 Hz). The minimum value is prone to be higher for [ʂ] (5,549-7,203 Hz) than [s] (3,212-4,721 Hz). Compared with the order for the native speaker (Figure 3c), it may be considered that Cantonese Female 1 cannot differentiate the retroflex [ʂ], the denti-alveolar [s] and alveolo-palatal [ɕ].

For Cantonese Female 1, the mean frequencies of three tokens of the noise peak and the minimum value of the noise range for [s] are 7,294 Hz and 4,055 Hz, which is similar to the noise peak of the native's [s] (7,882 Hz) and apparently lower than the minimum value of the native's [s] (6,387 Hz). The noise peak and the minimum value for [ɕ] are 5,016 Hz and 2,579 Hz, which are much lower than those of the native's [ɕ] (6,277 Hz and 4,884 Hz). Those for [ʂ] are 3,950 Hz and 3,147 Hz, which is similar to those of the native's [ʂ] (3,022 Hz and 1,711 Hz). Thus, it may be considered that Cantonese Female 1 merges [s] into [ɕ], and [ɕ] into [ʂ].

3.5 Cantonese Female 2

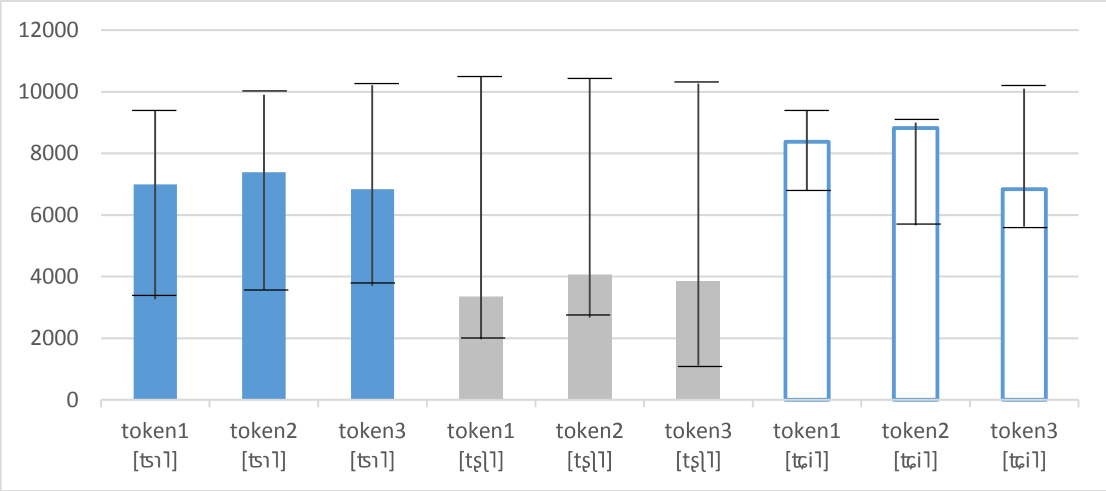


Figure 22. Frequencies in (Hz) of the noise peak (represented by a vertical bar) and noise range (represented by an 'I' line) for the Mandarin unaspirated affricates [ʦ, tʂ, ʨ] in the vowel context [i/ɿ/ʅ] produced by Cantonese Female 2.

Figure 22 shows the frequencies (in Hz) of the noise peak and the noise range for three tokens of each of the Mandarin unaspirated affricates [ʦ, tʂ, ʨ] in the vowel context [i/ɿ/ʅ] for Cantonese Female 2. Among the three sibilants, the noise peak is noticeably lower for the retroflex [tʂ] (3,362-4,079 Hz) than the denti-alveolar [ʦ] (6,835-7,386 Hz) and the alveolo-palatal [ʨ] (6,835-8,820 Hz). Between [ʦ] and [ʨ], the difference in the noise peak is not pronounced. As for the noise range, a noticeable difference among the three sibilants is in the minimum value of the noise range, which is also the lowest for [tʂ] (1,119-2,677 Hz). The minimum value of the noise range tends to be higher for [ʨ] (5,622-6,790 Hz) than [ʦ] (3,261-3,699 Hz). By comparison of the noise peak and noise range in a descending order of [ʦ] > [ʨ] > [tʂ] for the native speaker (Figure 3a), it may be considered that Cantonese Female 2 can produce the retroflex [tʂ], but cannot differentiate the denti-alveolar [ʦ] and alveolo-palatal [ʨ].

For Cantonese Female 2, the mean frequencies of three tokens of the noise peak and the minimum value of the noise range are 7,073 Hz and 3,512 Hz for [ʦ], similar to the noise peak of the native's [ʦ] (7,882 Hz) and apparently lower than the minimum value of the native's [ʦ] (6,387 Hz), and 8,011 Hz and 6,027 Hz for [ʨ], much higher than those of the native's [ʨ] (6,277 Hz and 4,884 Hz). Thus, it may suggest that Cantonese Female 2 merges the sibilants [ʦ] into [tʂ], and [ʨ] into [ʦ].

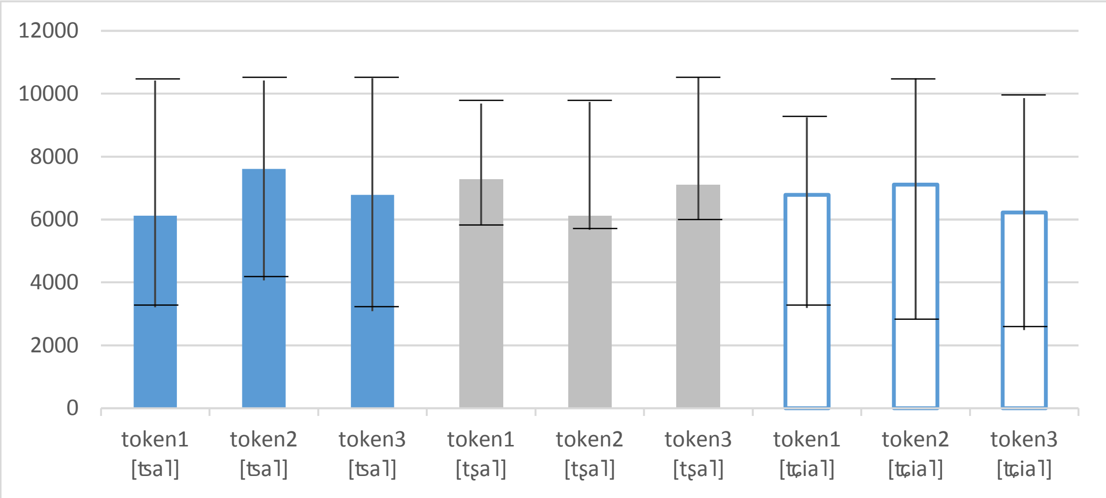


Figure 23. Frequencies in (Hz) of the noise peak (represented by a vertical bar) and noise range (represented by an 'I' line) for the Mandarin unaspirated affricates [ʦ, tʂ, ʨ] in the vowel context [a/ia] produced by Cantonese Female 2.

Figure 23 displays the frequencies of the noise peak and the noise range for three tokens of each of the Mandarin unaspirated affricates [ʦ, tʂ, ʨ] in the vowel context [a/ia] for Cantonese Female 2. Of the three sibilants, the noise peak for the denti-alveolar [ʦ] (6,118-7,607 Hz), the retroflex [tʂ] (6,118-7,276 Hz), and the alveolo-palatal [ʨ] (6,229-7,111 Hz). The differences in the noise peak almost do not exist between [ʦ], [ʨ], and [tʂ]. Regarding minimum value of the noise range, a difference should be taken note of among the three sibilants, which the lowest is also [ʨ] (2,482-3,188 Hz). The minimum value clings to be higher for [tʂ] (5,670-6,011 Hz) than [ʦ] (3,090-4,064 Hz). Compared with the order for the native speaker (Figure 3a), it may be said that Cantonese Female 2 cannot distinguish the denti-alveolar [ʦ] and alveolo-palatal [ʨ], and the retroflex [tʂ].

For Cantonese Female 2, 6,835 Hz and 3,455 Hz are the mean frequencies of three tokens of the noise peak and the minimum value of the noise range for [ʦ], apparently lower than those of the native's [ʦ] (7,882 Hz and 6,387 Hz). It is the opposite case for [tʂ] that the noise peak and the minimum value in [tʂ] are 6,835 Hz and 5,832 Hz, which is much higher than those of the native in [tʂ] (3,022 Hz and 1,711 Hz). Those of [ʨ] are 6,706 Hz and 2,839 Hz, which is similar to the noise peak of the native in [ʨ] (6,277 Hz), but distinctly lower is than the minimum value of the native's [ʨ] (4,884 Hz). Hence, it may suggest that Cantonese Female 2 merges the sibilants [ʦ] and [tʂ] into [ʨ], and [ʨ] into [tʂ].

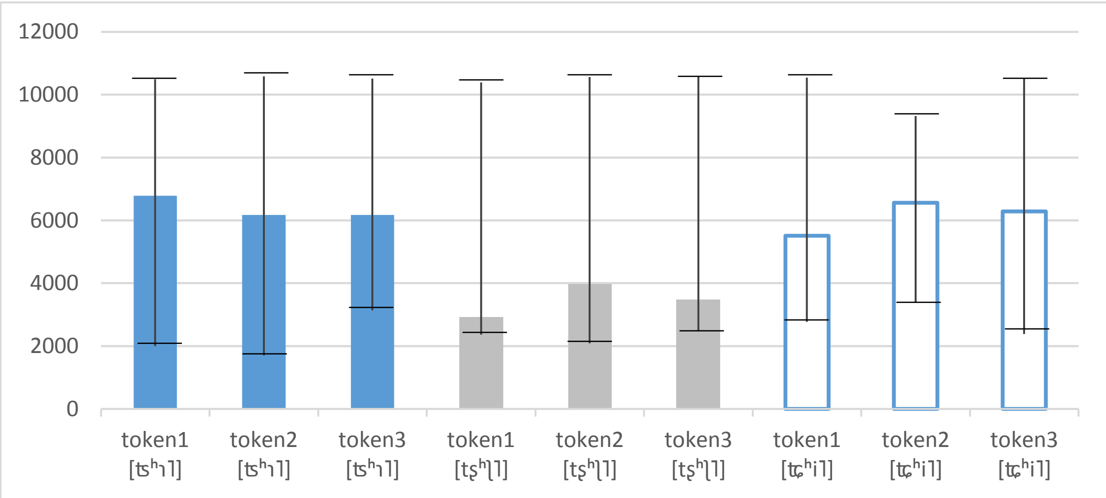


Figure 24. Frequencies in (Hz) of the noise peak (represented by a vertical bar) and noise range (represented by an 'I' line) for the Mandarin aspirated affricates [ʦʰ, tʂʰ, ʨʰ] in the vowel context [i/ɿ/ʅ] produced by Cantonese Female 2.

Figure 24 describes the frequencies of the noise peak and the noise range for three tokens of each of the Mandarin aspirated affricates [ʦʰ, tʂʰ, ʨʰ] in the vowel context [i/ɿ/ʅ] for Cantonese Female 2. Within the three sibilants, the noise peak is clearly lower for the retroflex [tʂʰ] (2,921-3,969 Hz), than the denti-alveolar [ʦʰ] (6,174-6,780 Hz) and the alveolo-palatal [ʨʰ] (5,512-6,559 Hz). The difference in the noise peak between [ʦʰ] and [ʨʰ] is not pronounced. Concerning the noise range, the three sets of sibilants in the minimum value of the noise range are very similar, which are 1,703-3,319 Hz for [ʦʰ], 2,093-2,482 Hz for [tʂʰ], 2,385-3,382 Hz for [ʨʰ]. By comparison of the noise peak and noise range in a descending order of [ʦʰ] > [ʨʰ] > [tʂʰ] for the native speaker (Figure 3b), it may be considered that Cantonese Female 2 can produce the retroflex [tʂʰ], but cannot differentiate the denti-alveolar [ʦʰ] and alveolo-palatal [ʨʰ].

For Cantonese Female 2, the mean frequencies of three tokens of the noise peak and the minimum value of the noise range for [ʦʰ] are 6,376 Hz and 2,279 Hz, apparently lower than those of the native's [ʦʰ] (7,882 Hz and 6,387 Hz), and 6,118Hz and 2,847 Hz for [ʨʰ] are similar to the noise peak of the native's [ʨʰ] (6,277 Hz) but lower than the minimum value of the native's [ʨʰ] (4,884 Hz). Therefore, it can be said that Cantonese Female 2 merges the sibilant [ʦʰ] into [ʨʰ] and [tʂʰ], and [ʨʰ] into [tʂʰ].

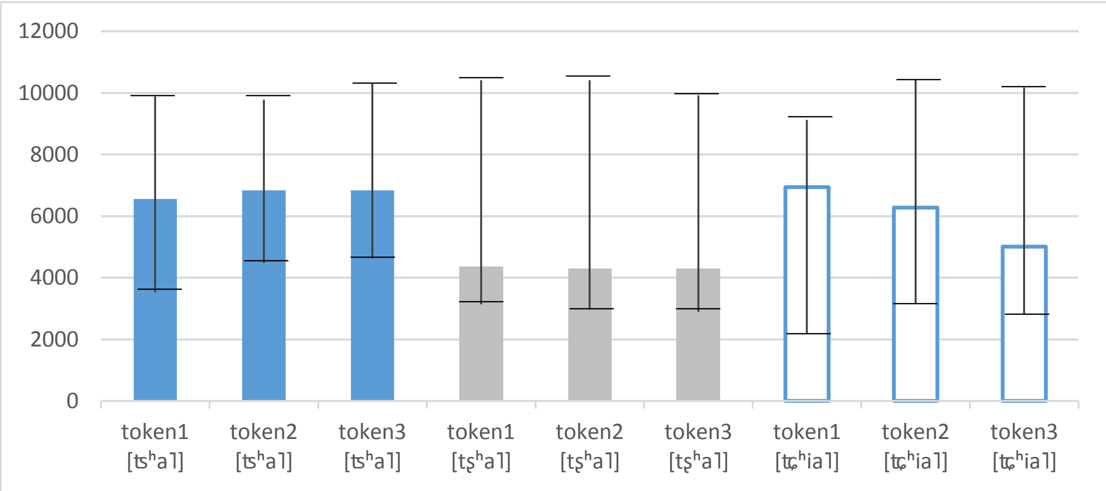


Figure 25. Frequencies in (Hz) of the noise peak (represented by a vertical bar) and noise range (represented by an 'I' line) for the Mandarin aspirated affricates [ʦʰ, tʂʰ, ʨʰ] in the vowel context [i/ɿ/ʅ] produced by Cantonese Female 2.

Figure 25 shows the frequencies of the noise peak and the noise range for three tokens of each of the Mandarin aspirated affricates [ʦʰ, tʂʰ, ʨʰ] in the vowel context [a/ia] for Cantonese Female 2. Among the three sibilants, the noise peak is apparently lower for the retroflex [tʂʰ] (4,299-4,354 Hz) than the alveolo-palatal [ʨʰ] (5,016-6,945 Hz) and the denti-alveolar [ʦʰ] (6,559-6835 Hz). The difference in the noise peak between [ʦʰ] and [ʨʰ] is not pronounced. As for the noise range, a relatively noticeable difference among the three sibilants is in the minimum value of the noise range, which is the highest for [ʦʰ] (3,528-4,624 Hz). The minimum value of the noise range tends to be lower for [tʂʰ] (2,896-3,139 Hz) than [ʨʰ] (2,190-3,188 Hz). Compared with the order for the native speaker (Figure 3b), it may be said that Cantonese Female 2 can differentiate the retroflex [tʂʰ] but not the denti-alveolar [ʦʰ] and alveolo-palatal [ʨʰ].

For Cantonese Female 2, the mean frequencies of three tokens of the noise peak and the minimum value of the noise range are 6,743 Hz and 4,210 Hz for [ʦʰ], obviously lower than those of the native's [ʦʰ] (7,882 Hz and 6,387 Hz). The noise peak and minimum value of [ʨʰ] are 6,081 Hz and 2,733 Hz, similar to the noise peak of the native's [ʨʰ] (6,277 Hz) but lower than the minimum value of the native's [ʨʰ] (4,884 Hz). Therefore, it may be suggest that Cantonese Female 2 here merges the sibilants [ʦʰ] into [ʨʰ], and [ʨʰ] into [tʂʰ].

Figure 26 shows the frequencies of the noise peak and the noise range for three tokens of each of the Mandarin fricatives [s, ʂ, ɕ] in the vowel context [i,ɿ,ʅ] for Cantonese Female 2. Of the three sibilants, the noise peak is apparently lower for the retroflex [ʂ] (3,858-3,969 Hz) than the denti-alveolar [s] (6,284-7,111 Hz) and the alveolo-palatal [ɕ] (5,953-6,890 Hz). The difference in the noise peak is not clear between [s] and [ɕ]. As for the noise range, the three sets of sibilants in the minimum value of the noise range are very similar, which are 2,993-3,675 Hz for [s], 2,239-2,798 Hz for [ʂ], and 2,263-2,628 Hz for [ɕ]. By comparison of the noise peak and noise range in a descending order of [s] > [ɕ] > [ʂ] for the native speaker (Figure 3c), it may be considered that Cantonese Female 2 can produce the retroflex [ʂ], but cannot distinguish the denti-alveolar [s] and alveolo-palatal [ɕ].

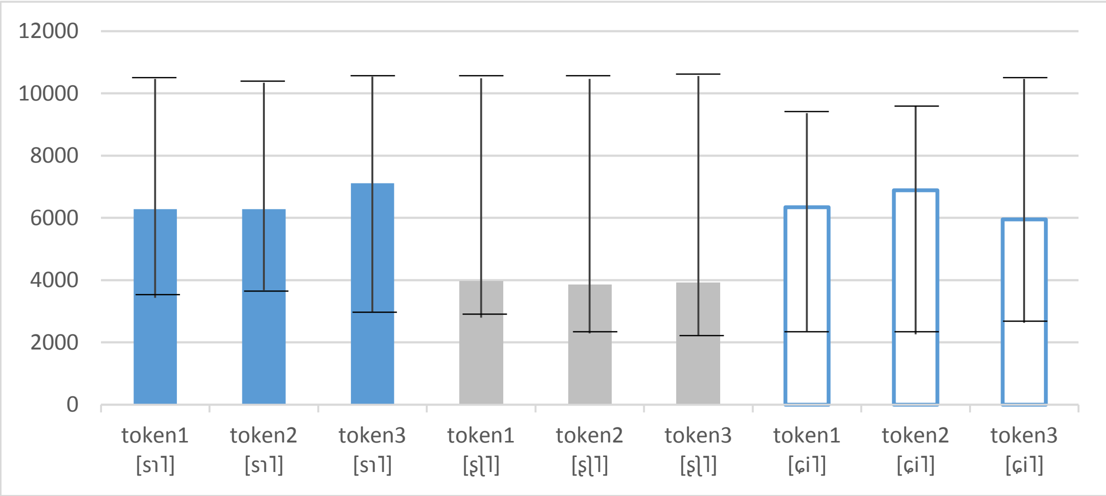


Figure 26. Frequencies in (Hz) of the noise peak (represented by a vertical bar) and noise range (represented by an 'I' line) for the Mandarin fricatives [s, ʂ, ɕ] in the vowel context [i/ɿ/ʅ] produced by Cantonese Female 2.

For Cantonese Female 2, the mean frequencies of three tokens of the noise peak and the minimum value of the noise range are 6,559 Hz and 3,366 Hz for [s], obviously lower than those of the native's [s] (7,882 Hz and 6,387 Hz). The noise peak and the minimum value for [ɕ] are 6,394 Hz and 2,417, similar to the noise peak of the native's [ɕ] (6,277 Hz) but lower than the minimum value of the native's [ɕ] (4,884 Hz). Thus, it may suggest that the Cantonese Female 2 merges the sibilants [s] into [ɕ], and [ɕ] into [ʂ].

Figure 27 reveals the frequencies of the noise peak and the noise range for three tokens of each of the Mandarin fricatives [s, ʂ, ɕ] in the vowel context [a/ia] for Cantonese Female 2. Within the three sibilants, the noise peak is apparently lower for the alveolo-palatal [ɕ] (4,630-6,174Hz) than the retroflex [ʂ] (7,221-7,386 Hz) and the denti-alveolar [s] (8,654 Hz), while there are two tokens of [s] (4,354-4,465 Hz) which are much lower. The difference between [s] and [ɕ] is not clear. Concerning the noise range, the three sets of sibilants in the minimum value of the noise range are very similar, which are 3,650-4,210 Hz for [ʂ], and 2,628-4,405 Hz for [ɕ], while 2,506-5,889 Hz for [s] are slightly higher. Compared with the order for the native speaker (Figure 3c), it may be considered that Cantonese Female 2 cannot differentiate the denti-alveolar [s], the retroflex [ʂ], and alveolo-palatal [ɕ].

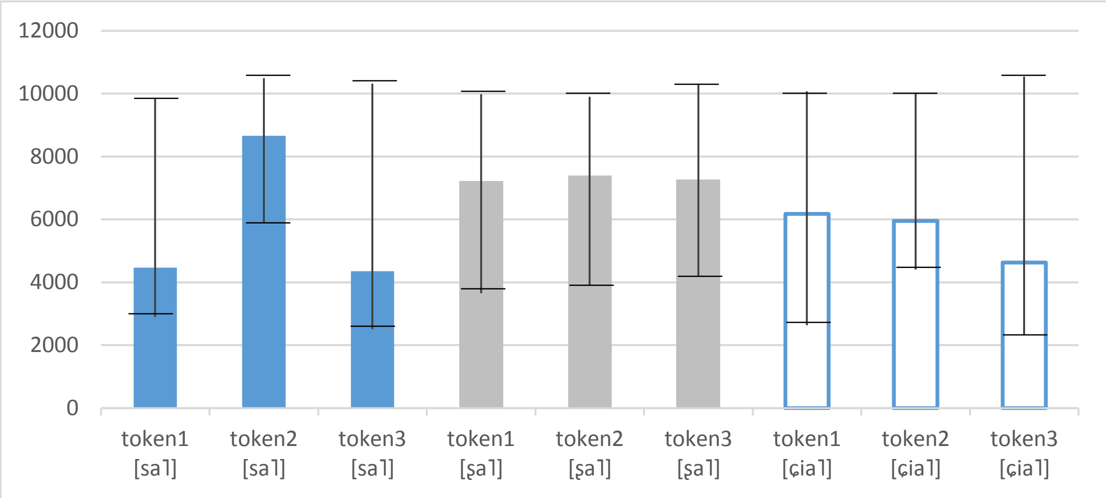


Figure 27. Frequencies in (Hz) of the noise peak (represented by a vertical bar) and noise range (represented by an 'I' line) for the Mandarin fricatives [s, ʂ, ɕ] in the vowel context [a/ia] produced by Cantonese Female 2.

For Cantonese Female 2, the mean frequencies of three tokens of the noise peak and the minimum value of the noise range for [s] are 5,824 Hz and 3,763 Hz, which are apparently lower than those of the native's [s] (7,882 Hz and 6,387 Hz), and 7,294 Hz and 3,926 Hz for [ʂ] are much higher than those of the native's [ʂ] (3,022Hz and 1,711 Hz). The noise peak and the minimum value for [ɕ] are 5,585 Hz and 3,123 Hz, similar to those of the native's [ɕ] (6,277 Hz and 4,884 Hz). Thus, it may be considered that Cantonese Female 2 merges [s] into [ɕ], and [ʂ] into [s] and [ɕ].

3.6 Summary of errors

After comparing the data of the 4 subjects with the native speaker, it is time to classify the types of errors according to the actual pronunciation of the subjects. Several possible patterns which are found in the data are listed as follows. (Details of each subject please refer to appendix 1). Letter 'X' refers to the target sibilant, and Letter 'Y' and 'Z' represent the non-target sibilant. 'New' is the sound which cannot be classified into any place categories, not produced within a boundary of any sets of sibilants, and/or not consistently mixed up with other sibilant equivalents. The formula to indicate the actual pronunciation is: 'target sibilant → [minimum value, peak value]'.

(1) Both the minimum and peak are **right**, which is placed into the category of distinguished sibilants.

X→[X, X]

(2) Both the minimum and peak are **wrong** as the same non-target sibilant, which is categorized into mispronunciation as other sibilants.

X→[Y/Z, Y/Z]

(3) Both the minimum and peak are **wrong** but they refer to two different sibilants respectively, placed into the category of inconsistency on mix-up with other sibilant equivalents.

X→[Y, Z] or X→[Z, Y]

(4) Either the minimum or peak is **wrong** but another part is **right**, also classified into the category of inconsistency on mix-up with other sibilant equivalents.

X→[X, Y/Z] or X→[Y/Z, X]

(5) As long as one parameter or both appear with mixed up sounds of two different sibilants, it is placed into new sound under the category of inconsistency on mix-up with other sibilant equivalents.

X→[X/Y/Z←New→Z/X/Y, X/Y/Z] or

X→[X/Y/Z, X/Y/Z←New→Z/X/Y] or

X→[X/Y/Z←New→Z/X/Y, X/Y/Z←New→Z/X/Y]

Because of our study aim which is to investigate whether the university students, the subjects, can distinguish the three sets of sibilants in production, therefore errors will be grouped according to the place of articulation of the target sibilant, facilitating the further discussion. As the total words pronounced of the 4 subjects are 216, thus this will be a good base for us to look at the distribution of the sibilants produced by the subjects in a big picture. Three big categories are: 1) Pattern type 1 is distinguished sibilants; 2) Pattern type 2 is mispronunciation as other sibilants; 3) Pattern type 3 to 5 are the category of inconsistency on mix-up with other sibilant equivalents.

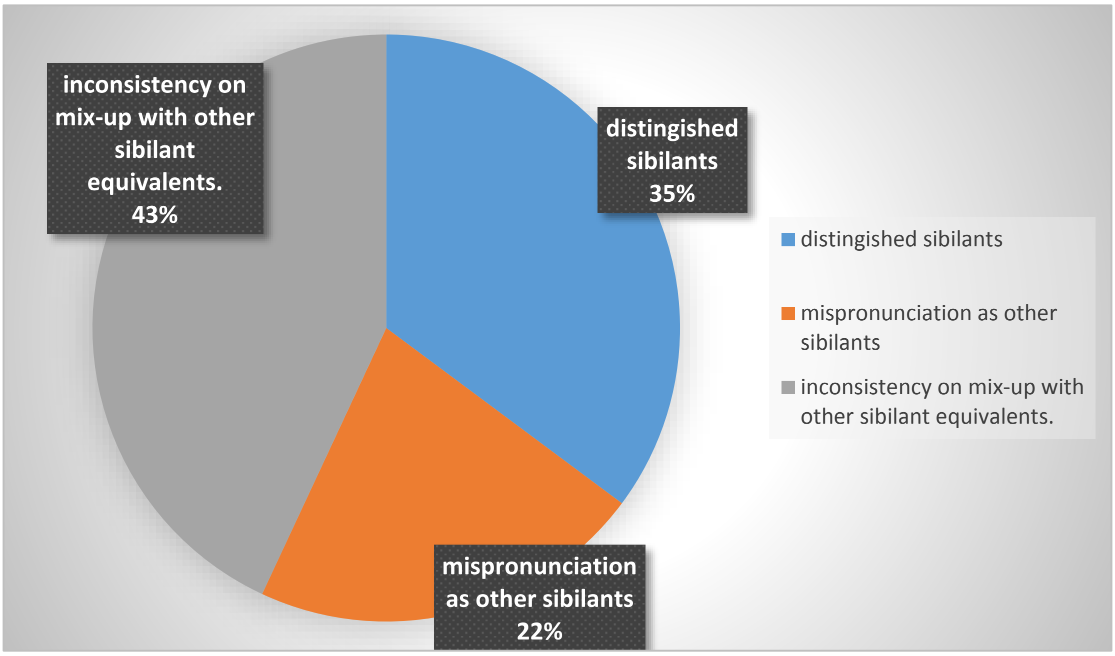


Figure 28. Distribution of the sibilants produced by the 4 subjects

Figure 28 shows that the subjects only can distinguish 35% of the sibilants, while 65% of them are mispronounced or inconsistently mixed up with others. This reveals that the subjects have not mastered the three sets of Mandarin sibilants very well.

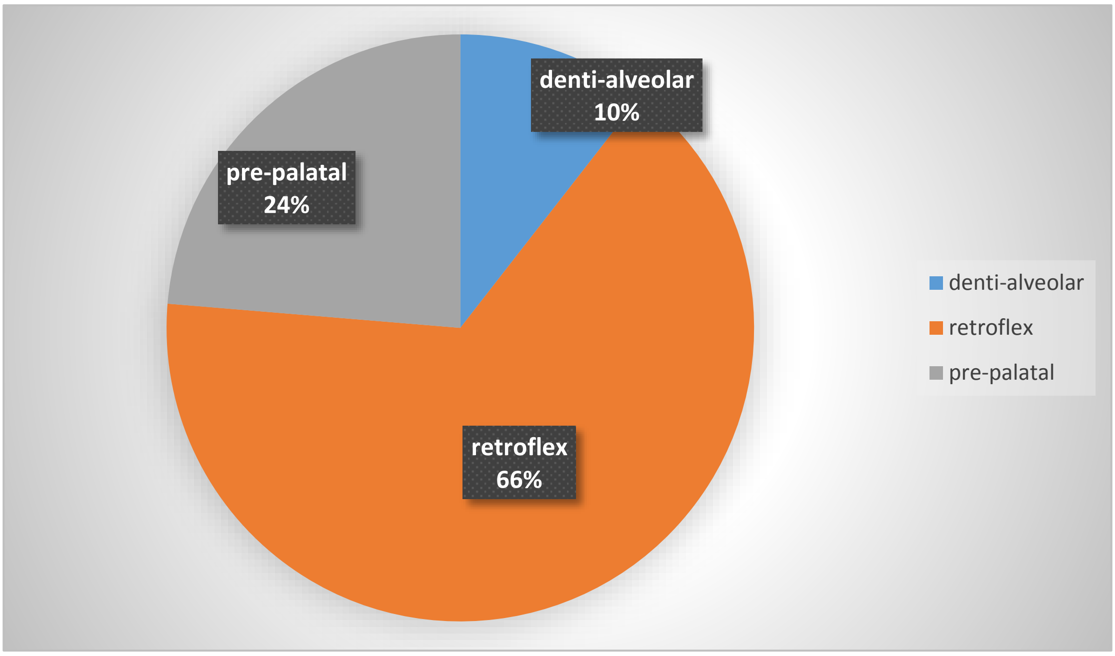


Figure 29. Distribution of distinguished sibilants

Figure 29 displays that among the three types of sibilants, retroflex sibilants are the group which the subjects mastered relatively well, compared with other two groups. The second one is alevolo-palatal and the last one is denti-alveolar group, showing that the subjects may have difficulties in learning or distinguishing them, leading to such smaller percentage.

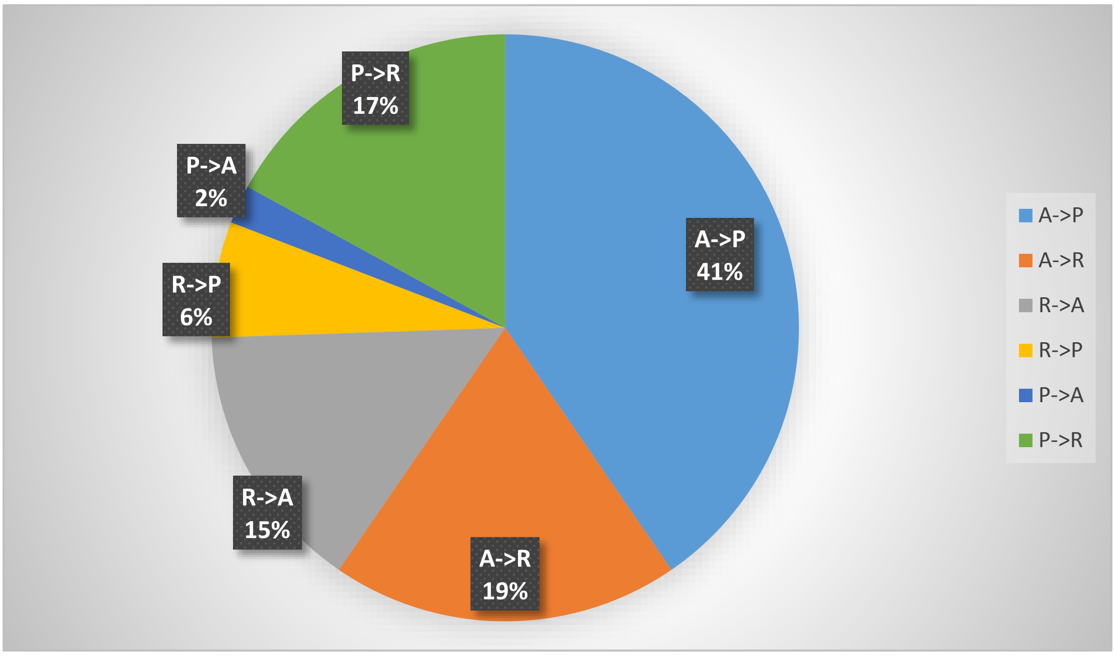


Figure 30. Distribution of mispronunciation as other sibilants

Apart from distinguished sibilants, the distribution of mispronunciation as other sibilants should be paid attention to as well. The alphabets have their meanings: A for denti-alveolar; R for retroflex; P for alevolo-palatal. The arrow stands for the change from one place of articulation to another one. The sibilants which are found to be mispronounced most is denti-alveolar to pre-palatal, 41%. The following types are denti-alveolar/alveolo-palatal to retroflex, which are about 20%, followed by retroflex altered to denti-alveolar, 15%. This means that for the subjects, denti-alveolar sibilants as the target sibilants are the most difficult one to be produced accurately among the three sets of sibilants. These changes indicate that the subjects considers some types of sibilants as the free variation.

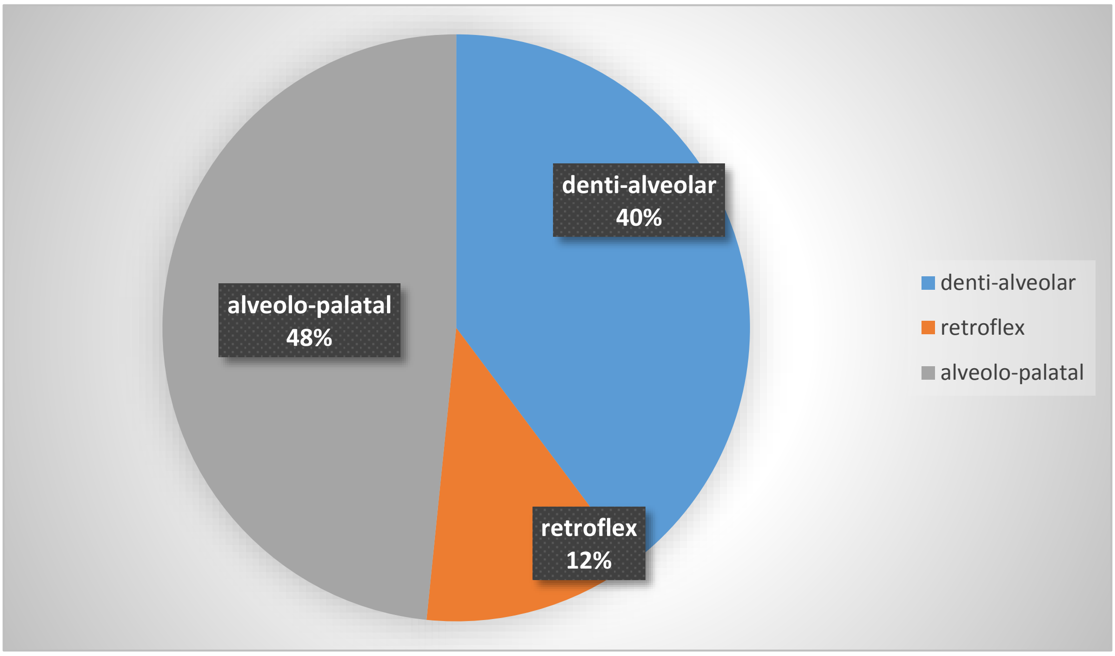


Figure 31. Inconsistency on mix-up with other sibilant equivalents.

Figure 31 describes the result of inconsistency on mix-up with other sibilant equivalents. Alveolo- palatal sibilants are the target sibilant which is the easiest to be mixed up with other sibilants, comprising of 48%. That means the subjects often consider alevolo-palatal sibilants as other sibilants. The similar situation also happens in denti-alveolar group, which comprises of 40%. The least one is retroflex group. It may be due to the reason which the retroflex group is more salient compared with others.

The patterns of inconsistently mixing up with other sibilants in denti-alveolar group is shown in figure 32. 60% are pronounced with features of alevolo-palatal sibilants, creating the pattern to be denti-alveolar altered to be half-alveolar and half-palatal in the peak and minimum value. The following pattern which is denti-alveolar changed to half-retroflex and half-pre-palatal contributes 24%. The third one is 11%, which new sound category appears if the target sibilants are denti-alveolar. The last one is denti-alveolar with retroflex characteristics. It can be said that the common problems which the subjects encounter are the denti-alveolar mixed up with alevolo-palatal. This explains why they pronounce some sounds in this category which seems not so accurate.

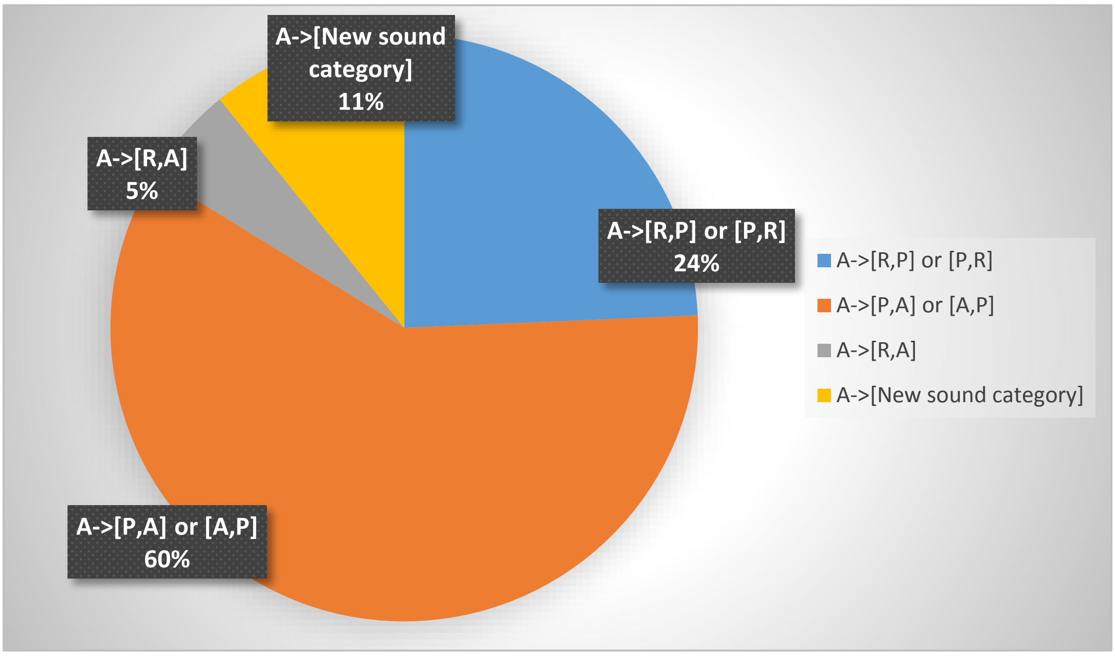


Figure 32. Distribution of inconsistency on mix-up with other sibilant equivalents for denti-alveolar sibilants.

The patterns of inconsistently mixing up with other sibilants in denti-alveolar group is shown in figure 32. 60% are pronounced with features of alevolo-palatal sibilants, creating the pattern to be denti-alveolar altered to be half-alveolar and half-palatal in the peak and minimum value. The following pattern which is denti-alveolar changed to half-retroflex and half-pre-palatal contributes 24%. The third one is 11%, which new sound category appears if the target sibilants are denti-alveolar. The last one is denti-alveolar with retroflex characteristics. It can be said that the common problems which the subjects encounter are the denti-alveolar mixed up with alevolo-palatal. This explains why they pronounce some sounds in this category which seems not so accurate.

In figure 33, the errors occurring in retroflex sibilants are displayed. 55% are the retroflex sibilants pronounced with the traits of both denti-alveolar and alveolo-palatal equivalents. 27% are similar to the former pattern, which consists the features of denti-alveolar and alveolo-palatal equivalents but the minimum value is placed in the boundary of retroflex. The last one which is 18% is mix-up with both denti-alveolar and alveolo-palatal sibilants. From this figure, it can be seen that retroflex sibilants are easily mixed up with denti-alveolar and alevolo-palatal equivalents, 82%, contributing to the inaccuracy of the pronunciation.

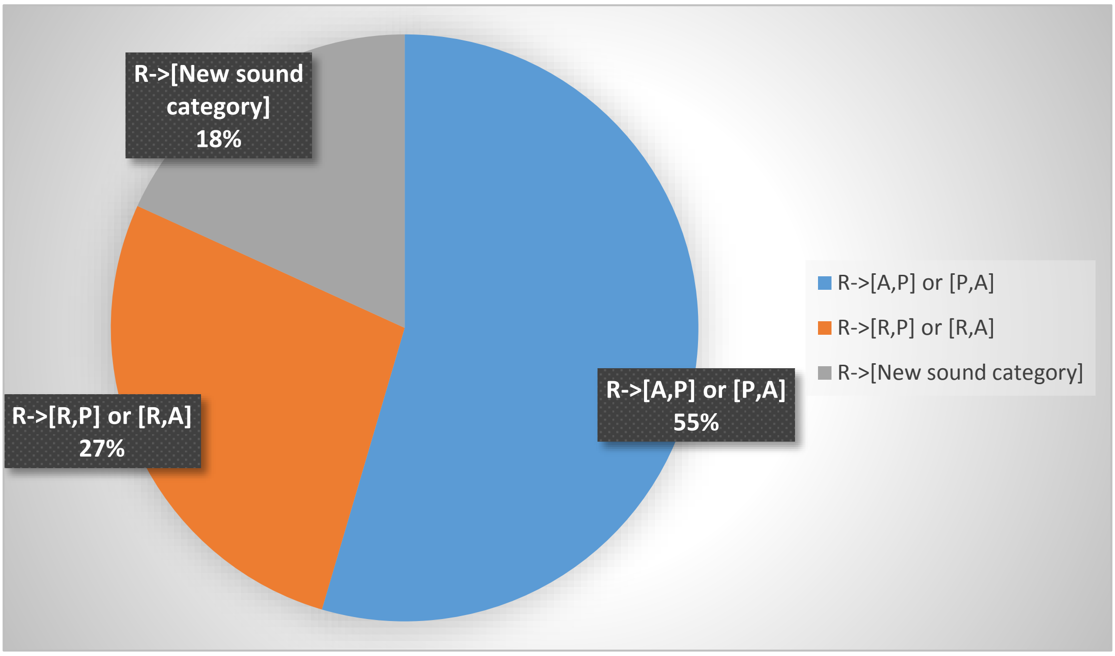


Figure 33. Distribution of inconsistency on mix-up with other sibilant equivalents for retroflex sibilants.

Figure 34 reveals that in alevolo-palatal sibilants the most serious problem is that the subjects pronounce the words with retroflex features, which is 89%. Other two types of errors consist of only about 10%. This explains that it is hard for the subjects to differentiate alveolo-palatal and retroflex sibilants.

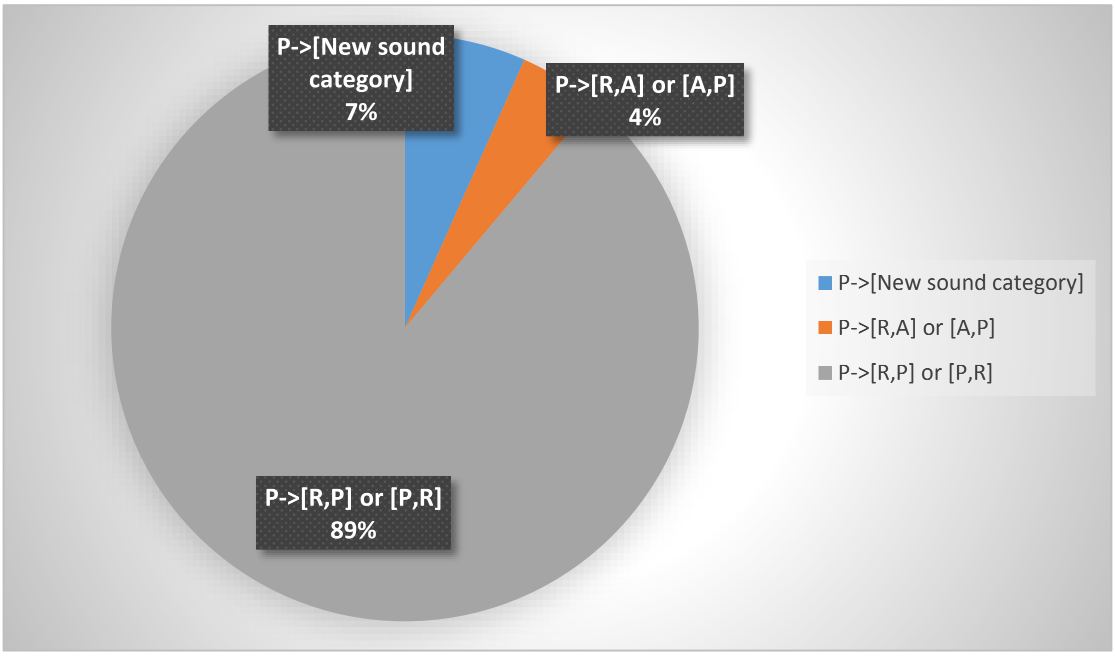


Figure 34. Distribution of inconsistency on mix-up with other sibilant equivalents for alevolo-palatal sibilants.

From the above detailed analysis, the subjects on average have not mastered the three sets of sibilants yet despite the 35% in distinguished sibilants which the performance in producing retroflex is the best. Their problems are classified into two types: 1) the target sibilant is mispronounced as the non-target sibilants, 2) inconsistency on mix-up with other sibilant equivalents for retroflex sibilants. These problems explain why the subjects are not able to distinguish the sibilants. The first type mainly roots in the mispronunciation that denti-alveolar sibilants are pronounced as alveolo-palatal or retroflex equivalents. The second type mainly appear in alevolo-palatal and denti-alveolar sibilants which are the target sibilant. For denti-alveolar, the problem is mix-up with alevolo-palatal sibilants; for retroflex, the problem is mix-up with denti-alveolar and alveolo-palatal sibilants; for alevolo-palatal, the problem is mix-up with retroflex sibilants.

4. Discussion

The above analysis of the data and the patterns of the errors contribute to and facilitate the following discussion with a focus on the place of articulation. This part is divided into three sections, which are 1) examine the analysis of the literature with the data, 2) compare the errors made by the subjects with those in another acoustic study, 3) try to investigate why the subjects may find the sibilants difficult to accurately produce with the speech learning model.

4.1 Examine the analysis of the literature with the data

Because of the reason that it is rare to see the acoustic analysis investigating the problems of Mandarin pronunciation of university students in Hong Kong, thus the results obtained can be compared with those coming from other approaches of analysis.

Tsang (1996) said that it is easy for Cantonese learners of Mandarin to confuse the three groups of sibilants, which denti-alveolar sibilants are the hardest while the other two types are not difficult. The result in our study shows that retroflex sibilants are the group with the highest percentage in distinguished sibilants, with the moderate level in mispronunciation, and with the lowest percentage in mix-up group. This matches the idea of Tsang, which retroflex sibilants are not the difficult one.

Furthermore, denti-alveolar is the hardest one due to the evidence of 10% in distinguished sibilants, 60% in mispronunciation, and 40% in mix-up. Overall, this type of sibilants is difficult in both differentiation and pronunciation, confirming his ideas again. However, it is not the case in alevolo-palatal sibilants because of 24% in distinguished sibilants and 48% in mix-up, despite 19% in mispronunciation. Although the percentage in mispronunciation is the lowest among other sibilants, still the percentage in distinguished sibilant is low and in mispronunciation is the highest. It means that alevolo-palatal sibilants are the second-hard one to be learnt but normally it is not easy to use other groups of sibilants to replace this set.

Ng (2001) thought that it is relatively hard for the Cantonese learners to produce the retroflex due to its absence in Cantonese. However, from our result, of three sets of sibilants, retroflex group is the one which has the highest percentage 66% in distinguished sibilants, almost 3 times of that in alveolo-palatal group, and 6 times of that in denti-alveolar group. This is opposite to the outcome from her analysis.

Lee-Wong (2013) believed that learners of Mandarin may have applied hypercorrection in retroflex due to the thought that there are so many retroflex sounds in Mandarin. This is true as from the statistic of last section non-retroflex sibilants are easily mispronounced as or mixed up with retroflex equivalents. The evidence is that 36% non-retroflex are mispronounced as retroflex in the category of mispronunciation, while 40% denti-alveolar and 96% in alveolo-palatal are mixed up with retroflex. This reveals that though retroflex sibilants are the learning target which is not difficult in pronunciation, still they are the one which is easily used to replace or be mixed up with others.

Hon (2003) proposed some error patterns in production of sibilants. The first type is confusion in unaspirated affricates which is considered as one of the most frequent errors. This is true, proved by the data obtained. Figure 35 displays the distribution of unaspirated affricates for the four Cantonese subjects. About 70% produced unaspirated affricates are placed into the category of either mispronunciation or mix-up, indicating that confusion in unaspirated affricates is common among Cantonese subjects.

The next pattern proposed is also proven by the data, which is retroflex fricative changed to palatal fricative. Figure 36 displays that 29 % of the retroflex fricatives are produced as alveolo-palatal equivalents or with the alveolo-palatal feature.

This matches Hon's idea again. The last pattern which aspirated alveolo-palatal affricates altered to aspirated denti-alveolar affricates is not found from the data.

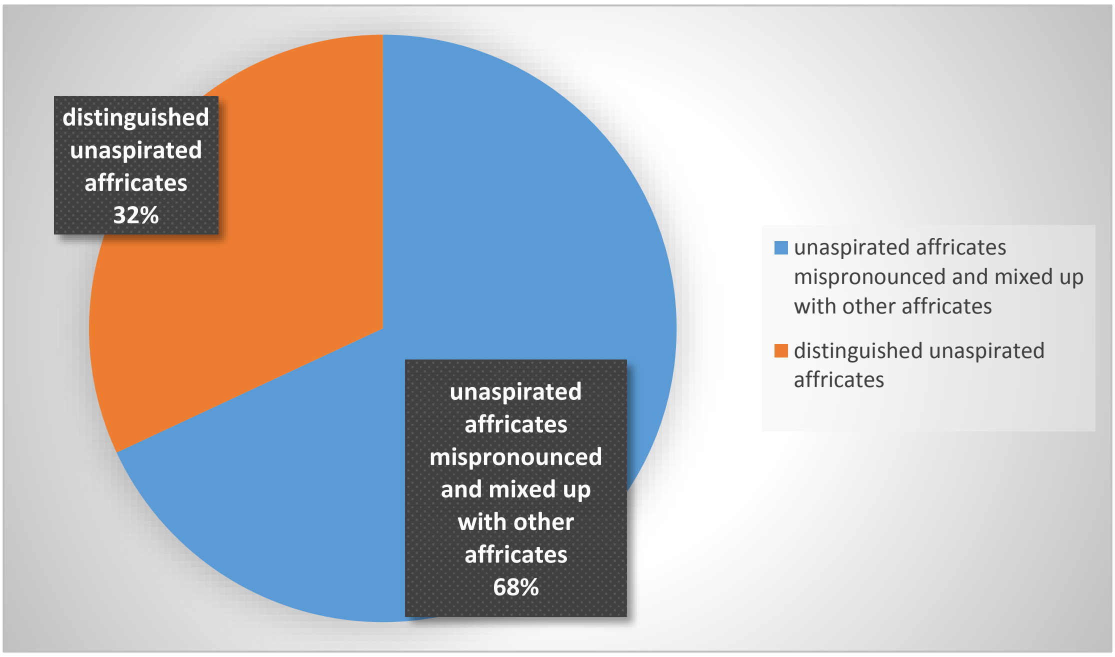


Figure 35. Distribution of unaspirated affricates for the four Cantonese subjects.

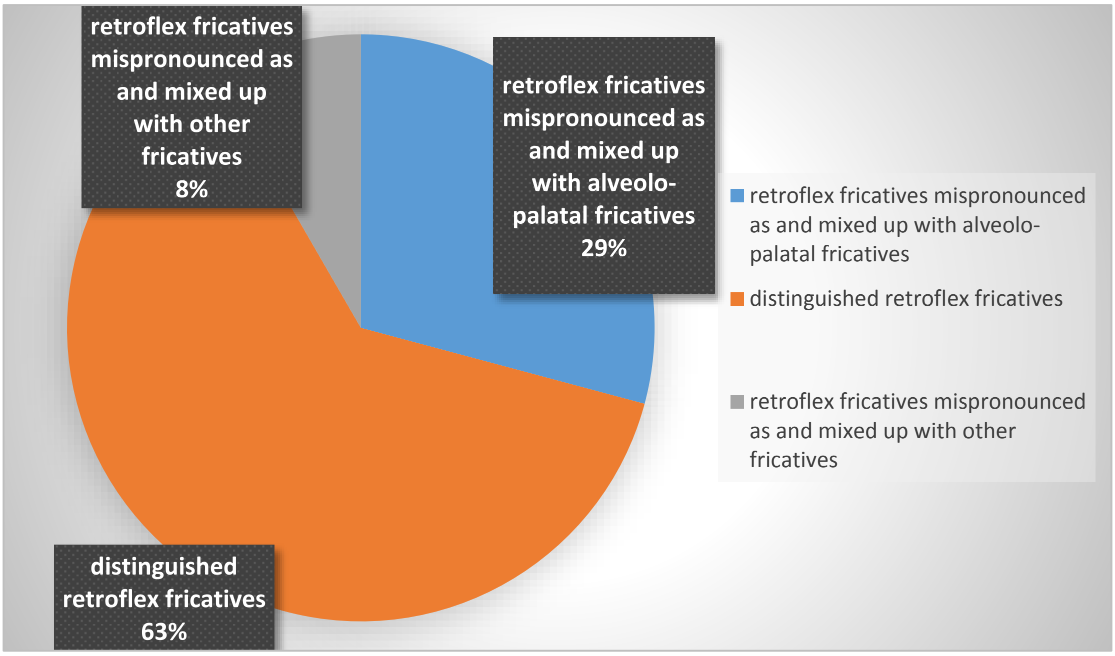


Figure 36. Distribution of retroflex fricatives for the four Cantonese subjects.

A lot of literature put emphasis on the negative transfer from Cantonese sibilants (Ng, 2001; Lee-Wong, 2013; Hon, 2003; Li, 2009; Wu and Su, 2014; Chung and Si, 2009). However, although this may be a foundational reason to explain why Cantonese speakers mispronounce or confuse with the sibilants, it is not workable in examining because this study does not have the acoustic analysis of the Cantonese sibilants. Therefore, if we do not have data about the Cantonese sibilants, it is hard to make comparison and examine the negative effect from Cantonese. However, the theoretical Cantonese influence is going to be discussed in section 4.3 with the speech learning model.

4.2 Compare the errors with those in the acoustic study by Chung and Si (2009)

The focus here will be the comparison of the errors produced by the subjects and the error patterns. It is worth doing such comparison because all the subjects in Chung and Si (2009)'s study and our study are the beginners of Mandarin. If they are in the similar level of proficiency in Mandarin, the errors and their patterns may be alike. The aim for this comparison is to investigate whether different learners of Mandarin from various places share similar errors or not. In the acoustic study of Chung and Si (2009), only the frequency of the peak is provided. Therefore, peaks will serve as the base for comparison.

4.2.1 Comparison with the subjects from different places

Table 10 describes the patterns of errors found in Chung and Si (2009)'s acoustic study. The patterns for Korean subjects are [s]or[ɕ]→[ʂ]; those for Japanese subjects are [s]→[ɕ], and [ɕ]→[ʂ]; those for Vietnamese subjects are [s]→[ʂ], and [s]or[ʂ]→[ɕ]. As the focus is the place of articulation, therefore the error patterns are rewritten in place categories as shown in table 11.

Country	Target sibilants	Actual pronunciation of the sibilants		
		[s]	[ʂ]	[ɕ]
Korea	[s]		✓	
	[ɕ]		✓	
Japan	[s]			✓
	[ɕ]		✓	
Vietnam	[s]		✓	✓
	[ʂ]			✓

Table 10. The distribution of the errors of Korean, Japanese, and Vietnamese subjects in the study of Chung and Si (2009).

Table 10 describes the patterns of errors found in Chung and Si (2009)'s acoustic study. The patterns for Korean subjects are [s]or[ɕ]→[ʂ]; those for Japanese subjects are [s]→[ɕ], and [ɕ]→[ʂ]; those for Vietnamese subjects are [s]→[ʂ], and [s]or[ʂ]→[ɕ]. As the focus is the place of articulation, therefore the error patterns are rewritten in place categories as shown in table 11.

Korean	**Japanese**	**Vietnamese**	**Cantonese**
A→R		***A→R***	***A→R***
	A→P	***A→P***	***A→P***
			R→ A
		R→P	***R→P***
			P→A
P→R	***P→R***		***P→R***

Table 11. The error patterns of the subjects from Korea, Japan, Vietnam, and Hong Kong.

Table 11 displays the patterns found in two studies. The patterns of the Cantonese learners of Mandarin come from figure 30 called the distribution of mispronunciation as other sibilants. It can be said that even though the learners come from different places, indeed most of them share the error patterns in terms of place category: A→R, A→P, R→P, and P→R (A: denti-alveolar; R: retroflex; P: alveolo-palatal). Besides, more patterns are displayed in the column of Cantonese that the Cantonese subjects have more problems compared with other learners.

4.3 Investigation with the speech learning model

Wu and Su (2014) use the speech learning model suggested by Flege (1995) to make prediction about which sibilants will be wrongly pronounced or not. The prediction is that /ʦ, ʦʰ, s/ are identical phones which will not be mispronounced, whereas /tʂ, tʂʰ, ʂ/ and /ʨ, ʨʰ, ɕ/ are new phones, which are easy to learn, leading to no errors. The result from their study may not be applicable for comparison because their subjects are advanced learners of Mandarin. Instead, this model can be the base of the part for further investigation.

Category	Denti-alveolar	Retroflex	Alveolo-palatal
Distinguished	10%	66%	24%
Mispronunciation	66%	21%	19%
Mix-up	40%	12%	48%

Table 12. The distribution of the sibilants in three categories for the four Cantonese subjects.

First, the laminal-alveolar sibilants in Cantonese and the denti-alveolar counterparts in Mandarin are not the identical consonants even though they share the manner of articulation and the property of the place category. Their actual pronunciations are different. Therefore, according to the model, denti-alveolar sibilants should be categorized as similar phones. This set of sibilants are the most difficult one for the learners, especially the beginners. This is supported by our data shown in table 12 that denti-alveolar group is the least distinguishable, the easiest to be mispronounced, and the second easy to be mixed up. The influence from Cantonese is that similar sounds in the Cantonese and Mandarin are one of the sources negatively affecting the learning of Mandarin.

Second, retroflex and alveolo-palatal sibilants are matched with the predication that they are relatively easier to be learnt. This is proven by the higher percentage in distinguished sibilants, and lower percentage in mix-up. However, in alveolo-palatal group, it is found that the percentage in mix-up is the highest. Therefore, it would be better to place it as the second difficult one.

The percentage of denti-alveolar and alveolo-palatal group in mix-up is very high. The reason may be that it is common for the Cantonese speakers to mix up with these two sets of sibilants and to find them similar in the sound quality. This implies that the concept of "similar" is not just about the manner or place of articulation, but also how similar the learners think of the sounds, from the perception perspective. From this angle, it can be classified as an intra-lingual error which the knowledge in one language affects each other.

The prediction is relatively not applicable to the beginners because it ignores that the learners need time to develop a new sound category in their mind and to adapt to the sounds at the target language. Hence, for beginners, it would be better to indicate the relative degree of difficulty of each sound at this stage. The degree of difficulty proposed in the beginning stage is that retroflex sibilants are the easiest; alveolo-palatal sibilants are the second-difficult; denti-alveolar sibilants are the most difficult.

5. Conclusion

This study aims to investigate the Mandarin sibilants produced by Cantonese speakers who are currently university students, through acoustic analysis. To facilitate the comparison, a native speaker of Mandarin who is currently an examiner of the National Putonghua Proficiency Test also participates in the sound recording. Frequency values of the noise peak and noise range are adopted to measure the sibilants produced by the subjects.

The differences in sound production between the subjects and the native are clearly analyzed in order to find out the patterns of errors. Five patterns, including one correct and four incorrect, are discovered from the Cantonese learners of Mandarin according to the distribution of frequency values for the noise peak and minimum of the noise range, listed in the following table. Letter 'X' is the target sibilants whereas other alphabets, such as 'Y' and 'Z' mean non-target sibilants, while the formula is 'target sibilant →[minimum, peak]'.

Types	Pattern	Description
1	X→[X, X]	Both the minimum and peak are **right**, which is placed into the category of distinguished sibilants.
2	X→[Y/Z, Y/Z]	Both the minimum and peak are **wrong** as the same non-target sibilant, which is categorized into mispronunciation as other sibilants.
3	X→[Y, Z] or X→[Z, Y]	Both the minimum and peak are **wrong** but they refer to two different sibilants respectively, placed into the category of

		inconsistency on mix-up with other sibilant equivalents.
4	X→[X,Y/Z] or X→[Y/Z, X]	Either the minimum or peak is **wrong** but another part is **right**, also classified into the category of inconsistency on mix-up with other sibilant equivalents.
5	X→[X/Y/Z←New→Z/X/Y, X/Y/Z] or X→[X/Y/Z, X/Y/Z←New→Z/X/Y] or X→[X/Y/Z←New→Z/X/Y, X/Y/Z←New→Z/X/Y]	As long as one parameter or both appear with mixed up sounds of two different sibilants, it is placed into new sound under the category of inconsistency on mix-up with other sibilant equivalents.

Table 13. Five types of production patterns for the four Cantonese subjects

Among all the test words pronounced, 35% of them are distinguished while 65% of them are mispronounced or mixed up with other sibilants. Of the distinguished sibilants, the subjects differentiate retroflex sibilants the best. For mispronunciation as other sibilants, denti-alveolar sibilants are the target which is the easiest to be mispronounced. In the mix-up category, both denti-alveolar and alveolo-palatal sibilants share the similar percentage, added to be about 90%. The detailed situation of mixing up in each place category is different. The denti-alveolar are easily mixed up with alevolo-palatal; retroflex with denti-alveolar and alveolo-palatal; alveolo-palatal with retroflex. All of these are the reference showing that the Cantonese subjects have not mastered the three sets of sibilants very well.

The result also examines the theoretical prediction of difficulty and the outcome generated by other methods. Denti-alveolar sibilants are the most difficult to be mastered. Retroflex sibilants often are the non-target sibilant which replaces the target sibilants. Retroflex sibilants altered to alveolo-palatal sibilants, and alveolo-palatal sibilants changed to denti-alveoalr sibilants are the patterns which exist in our data. The denti-alveolar sibilants in Mandarin and laminal alveolar counterparts in Cantonese are the similar phones, which are similar but not identical, increasing the difficulty of learning Mandarin. This negative effect comes from Cantonese. The degree of difficulty proposed in the beginning stage is retroflex > alveolo-palatal sibilants> denti-alveolar (from the easiest to the hardest).

Compared with the study conducted by Chung and Si (2009), it is found that some patterns of errors are shared among different learners of Mandarin, including the Cantonese subjects. They are: A→R, A→P, R→P, and P→R (A: denti-alveolar; R: retroflex; P: alveolo-palatal). It reveals that learners of Mandarin who come from various places may share some difficulties at the beginning stage.

To enhance the proficiency of the learners in the beginning stage, it is suggested that the Mandarin teachers can provide the minimal pair of the Mandarin sibilants as well as that of Mandarin and Cantonese sibilants to cultivate their ability to differentiate different sounds. The place of articulation of each sibilant can be explained in the lessons to let the learners understand what the differences between the sibilants are. The most important concept which should be taught is that in every language sounds carry meanings. This is the basic concept about language but it is rare for the beginners of a language understand it at the beginning stage. This idea will enhance the saliency towards the production of the three sets of sibilants in Mandarin.

Reference

Behrens, S. J., & Blumstein, S. E. (1988). Acoustic characteristics of English voiceless fricatives. Journal of Phonetics, 16, 295-298.

Cheng, Ting Au (鄭定歐). (1995).〈二論面向九七的香港推普工作〉,《一九九七與香港中國語文研討會論文集編輯委員會: 一九九七與香港中國語文研討會論文集》,香港: 香港中國語文學會.

Chung, Raung Fu (鍾榮富) and Si, Qiu Xue (司秋雪). (2009).〈從聲學分析探討初級韓日越學習者的華語擦音〉,《世界華語文教育學會：第九屆世界華語文教學研討會論文集第一冊語言分析(1)》, 台北: 世界華語文教育學會, 104-123.

Evers Vincent, Reetz Henning, and Lahiri Aditi. (1998). Crosslinguistic acoustic categorization of sibilants independent of phonological status. Journal of Phonetics, 26,345-370.

Flege, James E. (1995). Second language speech learning: Theories, findings, and problems. In W. Strange (Ed.), Speech Perception and Linguistic Experience: Issues in Cross-Language Research. Timonium: York Press, 233-272.

Hein, J. M. and Stevens, K. M. (1961). On the properties of voiceless fricative consonants. Journal of the Acoustical Society of America, 33, 589-596.

Hon, Wai Sun (韓維新). (2003).〈關於香港學生學習普通話發音難點的研究〉,《世界華語文教育學會: 第七屆世界華語文教學研討會論文集第六冊教學應用組》，台北: 世界華語文教育學會, 63-80.

Ng, Lai Ping (吳麗萍). (2001).〈如何在普通話語音學習上精益求精〉,《中國語文及人文素質教育》, 香港: 香港教育學院，185-197。

Nowak, P. M. (2006). The role of vowel transitions and frication noise in the perception of Polish sibilants. Journal of Phonetics, 34, 139-152.

Li, Fei (李斐). (2009).〈香港大學生普通話口語會話語音偏誤分析〉,《首都師範大學學報: 社會科學版》, S3, 80-86.

Ladefoged, P. and Johnson, K. (2011). A Course in Phonetics (6th ed.). Canada: Wadsworth, Cengage Learning, 11-16.

Lee-Wong Ping (李黃萍). (2013).〈語音偏誤分析—以香港學生為例〉,《深化語文教育改革的思考與實踐》, 南京: 東南大學出版社, 277-286.

Lee, W. S. and Zee, E. (2003). Standard Chinese (Beijing). Journal of the International Phonetic Association, 33(1), 109-112.

Tsang, Chi Fan (曾子凡). (1996).〈香港人學習普通話的難點〉,《方言》, 2, 110-112.

Wu, Chen Huei (吳貞慧) and Su, Jui Ya (蘇睿亞). (2014).〈澳門華語學習者華語語音偏誤類型之研究〉,《國立臺北教育大學語文集刊》, 26, 103-147.

Zee, E. (1999). Chinese (Hong Kong Cantonese). Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet. Cambridge: Cambridge University Press, 58-60.

Zhu, Xin Hua (祝新華), Chan, Shui Duen (陳瑞端), and Wen, Hong Bo (溫紅博). (2012).〈十年間香港大學生普通話水平的發展變化——基於香港理工大學的十年考試數據〉,《語言文字應用》, 1, 58-66.

<u>Appendixes:</u>

Appendix 1: Pronunciation of the Mandarin sibilants for Cantonese speakers

1.1 Cantonese Male 1

Target sibilant	Pronunciation in terms of the noise peak				Min. value of the noise range			
	[ʦ]	[tʂ]	[ʨ]	New	[ʦ]	[tʂ]	[ʨ]	New
[ʦ]	2		4		1		5	
[tʂ]		5		1		6		
[ʨ]			6		1	5		

Table 1. Pronunciation of the unaspirated affricates [ʦ, tʂ, ʨ]
(Total no. of token for each sibilant = 6)

Target sibilant	Pronunciation in terms of the noise peak				Min. value of the noise range			
	[ʦʰ]	[tʂʰ]	[ʨʰ]	New	[ʦʰ]	[tʂʰ]	[ʨʰ]	New
[ʦʰ]		3	3		1	3	2	
[tʂʰ]		6				6		
[ʨʰ]		4	2			4	2	

Table 2. Pronunciation of the aspirated affricates [ʦʰ, tʂʰ, ʨʰ]
(Total no. of token for each sibilant = 6)

Target sibilant	Pronunciation in terms of the noise peak				Min. value of the noise range			
	[s]	[ʂ]	[ɕ]	New	[s]	[ʂ]	[ɕ]	New
[s]	4	2			1		5	
[ʂ]		6				6		
[ɕ]		4	2			6		

Table 3. Pronunciation of the fricatives [s, ʂ, ɕ]
(Total no. of token for each sibilant = 6)

Type 1. X→[X, X] (Total: 20/54)		
Numbers	Details	Pattern
5	tʂ→[tʂ, tʂ]	R→[R, R]
1	ʦ→[ʦ, ʦ]	A→[A, A]
6	tʂʰ→[tʂʰ, tʂʰ]	R→[R, R]
1	ʨʰ→[ʨʰ, ʨʰ]	P→[P, P]
1	s→[s, s]	A→[A, A]
6	ʂ→[ʂ, ʂ]	R→[R, R]
Type 2. X→[Y, Y] (Total: 16/54)		
Numbers	Details	Pattern
4	ʦ→[ʨ, ʨ]	A→[P, P]
2	ʦʰ→[ʨʰ, ʨʰ]	A→[P, P]
3	ʦʰ→[tʂʰ, tʂʰ]	A→[R, R]
3	ʨʰ→[tʂʰ, tʂʰ]	P→[R, R]
4	ɕ→[ʂ, ʂ]	P→[R, R]
Type 3. X→[Y, Z] or X→[Z, Y] (Total: 2/54)		
Numbers	Details	Pattern
2	s→[ɕ, ʂ]	A→[P, R]
Type 4. X→[X, Y] or X→[Y, X] (Total: 15/54)		
Numbers	Details	Pattern
1	ʦ→[ʨ, ʦ]	A→[P, A]
5	ʨ→[tʂ, ʨ]	P→[R, P]
1	ʨ→[ʦ, ʨ]	P→[A, P]
1	ʦʰ→[ʦʰ, ʨʰ]	A→[A, P]
1	ʨʰ→[tʂʰ, ʨʰ]	P→[R, P]
1	ʨʰ→[ʨʰ, tʂʰ]	P→[P, R]
3	s→[ɕ, s]	A→[P, A]
2	ɕ→[ʂ, ɕ]	P→[R, P]
Type 5. X→[either peak or minimum value, or both is New category] (Total: 1/54)		
Numbers	Details	Pattern
1	tʂ→[tʂ, tʂ←→ ʨ]	R→[R, R← New →P]

Table 4. The production in five categories for Cantonese Male 1

1.2 Cantonese Male 2

Target sibilant	**Pronunciation in terms of the noise peak**				**Min. value of the noise range**			
	[ʦ]	[tʂ]	[ʨ]	New	[ʦ]	[tʂ]	[ʨ]	New
[ʦ]	4		2		1	2	3	
[tʂ]		5	1			5	1	
[ʨ]			6				6	

Table 5. Pronunciation of the unaspirated affricates [ʦ, tʂ, ʨ]
(Total no. of token for each sibilant = 6)

Target sibilant	**Pronunciation in terms of the noise peak**				**Min. value of the noise range**			
	[ʦ^h]	[tʂ^h]	[ʨ^h]	New	[ʦ^h]	[tʂ^h]	[ʨ^h]	New
[ʦ^h]		3	3			3	3	
[tʂ^h]		6				6		
[ʨ^h]			6			2	4	

Table 6. Pronunciation of the aspirated affricates [ʦ^h, tʂ^h, ʨ^h]
(Total no. of token for each sibilant = 6)

Target sibilant	**Pronunciation in terms of the noise peak**				**Min. value of the noise range**			
	[s]	[ʂ]	[ɕ]	New	[s]	[ʂ]	[ɕ]	New
[s]	6				1		5	
[ʂ]		4	2			4	2	
[ɕ]			6			1	5	

Table 7. Pronunciation of the fricatives [s, ʂ, ɕ]
(Total no. of token for each sibilant = 6)

Type 1. X→[X, X] (Total: 32/54)		
Numbers	Details	Pattern
1	ʦ→[ʦ, ʦ]	A→[A, A]
5	tʂ→[tʂ, tʂ]	R→[R, R]
6	ʨ→[ʨ, ʨ]	P→[P, P]
6	tʂʰ→[tʂʰ, tʂʰ]	R→[R, R]
4	ʨʰ→[ʨʰ, ʨʰ]	P→[P, P]
1	s→[s, s]	A→[A, A]
4	ʂ→[ʂ, ʂ]	R→[R, R]
5	ɕ→[ɕ, ɕ]	P→[P, P]
Type 2. X→[Y, Y] (Total: 9/54)		
Numbers	Details	Pattern
1	tʂ→[ʨ, ʨ]	R→[P, P]
3	ʦʰ→[ʨʰ, ʨʰ]	A→[P, P]
3	ʦʰ→[tʂʰ, tʂʰ]	A→[R, R]
2	ʂ→[ɕ, ɕ]	R→[P, P]
Type 3. X→[Y, Z] or X→[Z, Y] (Total: 2/54)		
Numbers	Details	Pattern
2	ʦ→[tʂ, ʨ]	A→[R, P]
Type 4. X→[X, Y] or X→[Y, X] (Total: 11/54)		
Numbers	Details	Pattern
3	ʦ→[ʨ, ʦ]	A→[P, A]
2	ʨʰ→[tʂʰ, ʨʰ]	P→[R, P]
5	s→[ɕ, s]	A→[P, A]
1	ɕ→[ʂ, ɕ]	P→[R, P]
Type 5. X→[either peak or minimum value, or both is New category] (Total: 0/54)		
none		

Table 8. The production in five categories for Cantonese Male 2

1.3 Cantonese Female 1

Target sibilant	**Pronunciation in terms of the noise peak**				**Min. value of the noise range**			
	[ʦ]	[tʂ]	[ʨ]	New	[ʦ]	[tʂ]	[ʨ]	New
[ʦ]	1		4	1	2		2	2
[tʂ]	3	1	2		3	2		1
[ʨ]		1	5			6		

Table 9. Pronunciation of the unaspirated affricates [ʦ, tʂ, ʨ]
(Total no. of token for each sibilant = 6)

Target sibilant	**Pronunciation in terms of the noise peak**				**Min. value of the noise range**			
	[ʦ^h]	[tʂ^h]	[ʨ^h]	New	[ʦ^h]	[tʂ^h]	[ʨ^h]	New
[ʦ^h]	3		3		1		4	1
[tʂ^h]	3	3				4	1	1
[ʨ^h]			6				6	

Table 10. Pronunciation of the aspirated affricates [ʦ^h, tʂ^h, ʨ^h]
(Total no. of token for each sibilant = 6)

Target sibilant	**Pronunciation in terms of the noise peak**				**Min. value of the noise range**			
	[s]	[ʂ]	[ɕ]	New	[s]	[ʂ]	[ɕ]	New
[s]	4	1	1		2	2	2	
[ʂ]	3	2	1		2	3	1	
[ɕ]			6			6		

Table 11. Pronunciation of the fricatives [s, ʂ, ɕ]
(Total no. of token for each sibilant = 6)

Type 1. X→[X, X] (Total: 9/54)		
Numbers	Details	Pattern
1	ʦ→[ʦ, ʦ]	A→[A, A]
1	tʂ→[tʂ, tʂ]	R→[R, R]
3	tʂʰ→[tʂʰ, tʂʰ]	R→[R, R]
2	s→[s, s]	A→[A, A]
2	ʂ→[ʂ, ʂ]	R→[R, R]
Type 2. X→[Y, Y] (Total: 11/54)		
Numbers	Details	Pattern
2	ʦ→[ʨ, ʨ]	A→[P, P]
3	tʂ→[ʦ, ʦ]	R→[A, A]
1	ʨ→[tʂ, tʂ]	P→[R, R]
1	ʦʰ→[ʨʰ, ʨʰ]	A→[P, P]
1	s→[ʂ, ʂ]	A→[R, R]
1	s→[ɕ, ɕ]	A→[P, P]
2	ʂ→[s, s]	R→[A, A]
Type 3. X→[Y, Z] or X→[Z, Y] (Total: 2/54)		
Numbers	Details	Pattern
1	tʂʰ→[ʨʰ, ʦʰ]	R→[P, A]
1	ʂ→[ɕ, s]	R→[P, A]
Type 4. X→[X, Y] or X→[Y, X] (Total: 27/54)		
Numbers	Details	Pattern
1	ʦ→[ʦ, ʨ]	A→[A, P]
1	tʂ→[tʂ, ʨ]	R→[R, P]
5	ʨ→[tʂ, ʨ]	P→[R, P]
3	ʦʰ→[ʨʰ, ʦʰ]	A→[P, A]
1	ʦʰ→[ʦʰ, ʨʰ]	A→[A, P]
1	tʂʰ→[tʂʰ, ʦʰ]	R→[R, A]
6	ʨʰ→[tʂʰ, ʨʰ]	P→[R, P]
1	ʂ→[ʂ, ɕ]	R→[R, P]
1	s→[ʂ, s]	A→[R, A]
1	s→[ɕ, s]	A→[P, A]
6	ɕ→[ʂ, ɕ]	P→[R, P]
Type 5. X→[either peak or minimum value, or both is New category] (Total: 5/54)		
Numbers	Details	Pattern

1	ʦ→[tʂ←→ʨ, tʂ←→ʨ]	A→[R←New→P, R←New →P]
1	ʦ→[tʂ←→ʨ, ʨ]	A→[R← New →P, P]
1	tʂ→[tʂ←→ ʨ, ʨ]	R→[R← New →P, P]
1	ʦʰ→[ʨʰ←→ ʦʰ, ʨʰ]	A→[P← New →A, P]
1	tʂʰ→[ʨʰ←→ʦʰ, ʦʰ]	R→[P← New →A, A]

Table 12. The production in five categories for Cantonese Female 1

1.4 Cantonese Female 2

Target sibilant	**Pronunciation in terms of the noise peak**				**Min. value of the noise range**			
	[ʦ]	[tʂ]	[ʨ]	New	[ʦ]	[tʂ]	[ʨ]	New
[ʦ]	3		3			2	4	
[tʂ]	2	3	1		3	3		
[ʨ]	3		3		1	3		2

Table 13. Pronunciation of the unaspirated affricates [ʦ, tʂ, ʨ]
(Total no. of token for each sibilant = 6)

Target sibilant	**Pronunciation in terms of the noise peak**				**Min. value of the noise range**			
	[ʦʰ]	[tʂʰ]	[ʨʰ]	New	[ʦʰ]	[tʂʰ]	[ʨʰ]	New
[ʦʰ]			6			3	3	
[tʂʰ]		6				6		
[ʨʰ]			6			5	1	

Table 14. Pronunciation of the aspirated affricates [ʦʰ, tʂʰ, ʨʰ]
(Total no. of token for each sibilant = 6)

Target sibilant	**Pronunciation in terms of the noise peak**				**Min. value of the noise range**			
	[s]	[ʂ]	[ɕ]	New	[s]	[ʂ]	[ɕ]	New
[s]	2	2	2		1	3	2	
[ʂ]	3	3				3	3	
[ɕ]			5	1		5	1	

Table 15. Pronunciation of the fricatives [s, ʂ, ɕ]
(Total no. of token for each sibilant = 6)

Type 1. X→[X, X] (Total: 15/54)		
Numbers	Details	Pattern
3	tʂ → [tʂ, tʂ]	R→[R, R]
6	tʂʰ → [tʂʰ, tʂʰ]	R→[R, R]
1	ʨʰ → [ʨʰ, ʨʰ]	P→[P, P]
1	s → [s, s]	A→[A, A]
3	ʂ → [ʂ, ʂ]	R→[R, R]
1	ɕ → [ɕ, ɕ]	P→[P, P]
Type 2. X→[Y, Y] (Total: 11/54)		
Numbers	Details	Pattern
1	ʦ → [ʨ, ʨ]	A→[P, P]
2	tʂ → [ʦ, ʦ]	R→[A, A]
1	ʨ → [ʦ, ʦ]	P→[A, A]
3	tʂʰ → [ʨʰ, ʨʰ]	A→[P, P]
2	s → [ɕ, ɕ]	A→[P, P]
2	s → [ʂ, ʂ]	A→[R, R]
Type 3. X→[Y, Z] or X→[Z, Y] (Total: 10/54)		
Numbers	Details	Pattern
2	ʦ → [tʂ, ʨ]	A→[R, P]
1	tʂ → [ʦ, ʨ]	R→[A, P]
1	ʨ → [tʂ, ʦ]	P→[R, A]
3	ʦʰ → [tʂʰ, ʨ]	A→[R, P]
3	ʂ → [ɕ, s]	R→[P, A]
Type 4. X→[X, Y] or X→[Y, X] (Total: 15/54)		
Numbers	Details	Pattern
3	ʦ → [ʨ, ʦ]	A→[P, A]
2	ʨ → [tʂ, ʨ]	P→[R, P]
5	ʨʰ → [tʂʰ, ʨʰ]	P→[R, P]
1	s → [ʂ, s]	A→[R, A]
4	ɕ → [ʂ, ɕ]	P→[R, P]
Type 5. X→[either peak or minimum value, or both is New category] (Total: 3/54)		
Numbers	Details	Pattern
1	ʨ → [ʦ←→ ʨ, ʨ]	P→[A←→P, P]
1	ʨ → [ʦ←→ ʨ, ʦ]	P→[A←→P, A]
1	ɕ → [ʂ, ɕ←→ ʂ]	P→[R, P←→R]

Table 16. The production in five categories for Cantonese Female 2

Appendix 2: Frequency values of the noise range and noise peak

2.1 Native speaker

Target sibilant	Test words	Token no.	Noise range		Noise peak
			Min value	Max value	
[ʦ]	[ʦɿ˥] 資 'capital'	1	6566	10608	8544
		2	5597	10538	7221
		3	6027	10461	7331
	[ʦa˥] 紮 'tie'	1	8177	10440	8764
		2	6644	10416	7882
		3	7861	10392	8764
[ʦʰ]	[ʦʰɿ˥] 疵 'flaw'	1	5816	10586	7441
		2	4721	10416	7386
		3	5914	10513	8764
	[ʦʰa˥] 擦 'rub'	1	7909	10513	8544
		2	6011	10440	7384
		3	6376	10513	6835
[s]	[sɿ˥] 司 'in charge of'	1	5743	10659	7497
		2	6522	10562	7111
		3	5646	10611	7717
	[sa˥] 撒 'cast'	1	6352	10440	8875
		2	5962	10562	7938
		3	7130	10538	7882
[tʂ]	[tʂʅ˥] 知 'know'	1	1508	8834	3969
		2	1654	9150	2315
		3	1922	9370	3528
	[tʂa˥] 渣 'residue'	1	1727	9418	3528
		2	2239	8785	3913
		3	1800	9297	3803
[tʂʰ]	[tʂʰʅ˥] 吃 'eat'	1	1995	9978	3031
		2	1679	10294	2701
		3	1581	9929	3748
	[tʂʰa˥] 叉 'folk'	1	937	9467	2646
		2	924	7763	1929
		3	1557	7617	2646

Target sibilant	Test words	Token no.	Noise range		Noise peak
			Min value	Max value	
[ʂ]	[ʂʅ˥] 失 ‘lose’	1	1886	9922	2535
		2	1715	9751	2701
		3	1715	10143	2535
	[ʂa˥] 沙 ‘sand’	1	1703	8956	2480
		2	2409	9589	3252
		3	1849	10051	3142
[tɕ]	[tɕi˥] 基 ‘base’	1	5695	10412	7552
		2	6571	10314	7552
		3	5622	10513	7441
	[tɕia˥] 加 ‘add’	1	5719	10416	6229
		2	5329	10319	6559
		3	5573	10294	6615
[tɕʰ]	[tɕʰi˥] 七 ‘seven’	1	5743	10538	7552
		2	5135	10440	6229
		3	5938	10513	7000
	[tɕʰia˥] 掐 ‘nip off’	1	3942	10513	4740
		2	3188	10586	4354
		3	3017	10489	4795
[ɕ]	[ɕi˥] 希 ‘hope’	1	5695	10586	6284
		2	4064	10051	4740
		3	5597	10440	6615
	[ɕia˥] 蝦 ‘prawn’	1	3845	10586	6063
		2	4015	10319	5843
		3	3236	10416	6835

2.2 Cantonese Male 1

Target sibilant	Test words	Token no.	Noise range		Noise peak
			Min value	Max value	
[ʦ]	[ʦɿ˥] 資 ‘capital’	1	3942	10465	6559
		2	5597	10173	6890
		3	4916	10465	6559
	[ʦa˥] 紮 ‘tie’	1	4283	10075	8434
		2	6011	10489	7276
		3	3845	9540	6174
[ʦʰ]	[ʦʰɿ˥] 疵 ‘flaw’	1	4015	10075	5126
		2	5719	10465	6615
		3	4672	10465	6780
	[ʦʰa˥] 擦 ‘rub’	1	1630	7739	3969
		2	2531	7569	3031
		3	1947	10148	3583
[s]	[sɿ˥] 司 ‘in charge of’	1	5622	10586	7000
		2	5086	10538	6229
		3	5281	10367	6449
	[sa˥] 撒 ‘cast’	1	4259	10465	6615
		2	3650	10027	4410
		3	3699	10319	4134
[tʂ]	[tʂʅ˥] 知 ‘know’	1	1557	7861	2756
		2	1679	7593	2590
		3	1533	9150	2701
	[tʂa˥] 渣 ‘residue’	1	2701	8810	5126
		2	2433	8591	3472
		3	1654	8931	3417
[tʂʰ]	[tʂʰʅ˥] 吃 ‘eat’	1	1460	7690	2535
		2	1508	8372	2756
		3	1606	8712	2701
	[tʂʰa˥] 叉 ‘folk’	1	1752	7252	2866
		2	1581	8810	2866
		3	1679	8858	2756

Target sibilant	Test words	Token no.	Noise range		Noise peak
			Min value	Max value	
[ʂ]	[ʂʅ˥] 失 'lose'	1	1630	7666	2590
		2	1606	8007	3252
		3	1654	7544	2701
	[ʂa˥] 沙 'sand'	1	1727	8007	2866
		2	2604	7642	3197
		3	1679	7885	3528
[tɕ]	[tɕi˥] 基 'base'	1	2896	9345	6394
		2	5695	10465	7607
		3	2823	10538	6559
	[tɕia˥] 加 'add'	1	2823	9102	5512
		2	2750	9735	5677
		3	2701	9856	6559
[tɕʰ]	[tɕʰi˥] 七 'seven'	1	2896	10343	4575
		2	5622	10465	6615
		3	5037	10586	6615
	[tɕʰia˥] 掐 'nip off'	1	2823	9126	4410
		2	3748	9783	4244
		3	2677	9467	4244
[ɕ]	[ɕi˥] 希 'hope'	1	2579	10440	3087
		2	2847	10367	4410
		3	2871	10489	4189
	[ɕia˥] 蝦 'prawn'	1	2944	9686	3087
		2	2701	10416	3031
		3	3017	10294	3362

2.3 Cantonese Male 2

Target sibilant	Test words	Token no.	Noise range		Noise peak
			Min value	Max value	
[ʦ]	[ʦɿ˥] 資 ‘capital’	1	3675	10246	5788
		2	5622	10489	8489
		3	5256	10440	7000
	[ʦa˥] 紮 ‘tie’	1	3017	9491	6725
		2	4405	10538	7497
		3	4794	9881	7497
[ʦʰ]	[ʦʰɿ˥] 疵 ‘flaw’	1	4259	9832	6725
		2	4186	10294	6670
		3	5500	10465	5898
	[ʦʰa˥] 擦 ‘rub’	1	2093	8956	3528
		2	1874	9589	3252
		3	2044	8639	2976
[s]	[sɿ˥] 司 ‘in charge of’	1	6035	10392	8544
		2	4502	10465	7386
		3	4624	10416	7662
	[sa˥] 撒 ‘cast’	1	4672	10246	7441
		2	5573	10416	7552
		3	5378	9662	7111
[tʂ]	[tʂʅ˥] 知 ‘know’	1	3553	9321	6504
		2	2336	9248	3583
		3	2263	9321	4354
	[tʂa˥] 渣 ‘residue’	1	1849	7617	3197
		2	2141	7203	3031
		3	2287	7179	3142
[tʂʰ]	[tʂʰʅ˥] 吃 ‘eat’	1	2506	8883	4410
		2	2117	9029	3583
		3	2531	7569	3417
	[tʂʰa˥] 叉 ‘folk’	1	2798	9272	3197
		2	2093	9345	2535
		3	2409	7617	3528

Target sibilant	Test words	Token no.	Noise range		Noise peak
			Min value	Max value	
[ʂ]	[ʂʅ˥] 失 'lose'	1	2677	9150	4630
		2	2823	9735	3417
		3	4478	10440	5843
	[ʂa˥] 沙 'sand'	1	3090	9102	3472
		2	2798	7690	3472
		3	3553	8858	4906
[tɕ]	[tɕi˥] 基 'base'	1	3699	9491	6559
		2	4259	8907	5622
		3	4113	9491	5677
	[tɕia˥] 加 'add'	1	3455	9564	6449
		2	3504	9662	5898
		3	3675	9297	5622
[tɕʰ]	[tɕʰi˥] 七 'seven'	1	5500	9491	6780
		2	4770	10197	5898
		3	4551	10270	4961
	[tɕʰia˥] 掐 'nip off'	1	2555	9954	4630
		2	3528	9540	5788
		3	2214	10367	4575
[ɕ]	[ɕi˥] 希 'hope'	1	4648	9540	6615
		2	4405	10124	6615
		3	4526	9978	6835
	[ɕia˥] 蝦 'prawn'	1	2093	10489	4795
		2	4210	10246	5898
		3	4575	8980	6945

2.4 Cantonese Female 1

Target sibilant	Test words	Token no.	Noise range		Noise peak
			Min value	Max value	
[ʦ]	[ʦɿ˥] 資 ‘capital’	1	3309	10270	5843
		2	3382	9516	4685
		3	4113	10611	6504
	[ʦa˥] 紮 ‘tie’	1	5037	10586	6063
		2	5889	10611	8544
		3	5865	10513	6229
[ʦʰ]	[ʦʰɿ˥] 疵 ‘flaw’	1	4794	10538	7056
		2	5719	10611	6339
		3	5500	10635	7111
	[ʦʰa˥] 擦 ‘rub’	1	5573	10489	6118
		2	4745	10513	8379
		3	4843	10659	6890
[s]	[sɿ˥] 司 ‘in charge of’	1	2628	10148	3858
		2	5987	10440	7221
		3	5743	10489	7276
	[sa˥] 撒 ‘cast’	1	3212	10684	7386
		2	4721	10611	8268
		3	4234	10586	6229
[tʂ]	[tʂʅ˥] 知 ‘know’	1	2920	9954	3913
		2	3212	10270	6559
		3	3334	10562	6780
	[tʂa˥] 渣 ‘residue’	1	6060	10513	7221
		2	5865	10586	7552
		3	5816	10659	7056
[tʂʰ]	[tʂʰʅ˥] 吃 ‘eat’	1	2969	9954	3803
		2	2385	9077	3693
		3	2312	9126	3472
	[tʂʰa˥] 叉 ‘folk’	1	3455	10586	7111
		2	4478	10611	7056
		3	5670	10562	7000

Target sibilant	Test words	Token no.	Noise range		Noise peak
			Min value	Max value	
[ʂ]	[ʂʅ˥] 失 ‘lose’	1	2068	10173	3472
		2	2020	10246	5016
		3	2020	10124	3362
	[ʂa˥] 沙 ‘sand’	1	5549	10538	7331
		2	7203	10586	8820
		3	6327	10416	7056
[tɕ]	[tɕi˥] 基 ‘base’	1	2385	10513	4851
		2	2506	10416	6174
		3	2579	10270	5953
	[tɕia˥] 加 ‘add’	1	2239	9856	3528
		2	2579	10075	5016
		3	2458	9321	4906
[tɕʰ]	[tɕʰi˥] 七 ‘seven’	1	3261	10513	6504
		2	2214	10562	4961
		3	2652	10611	6008
	[tɕʰia˥] 掐 ‘nip off’	1	2652	9856	5016
		2	2823	9516	5016
		3	2287	10319	4961
[ɕ]	[ɕi˥] 希 ‘hope’	1	2506	10538	4906
		2	3090	10440	6229
		3	3285	10416	6670
	[ɕia˥] 蝦 ‘prawn’	1	2579	10246	5016
		2	2458	10294	5016
		3	2701	10416	5016

2.5 Cantonese Female 2

Target sibilant	Test words	Token no.	Noise range		Noise peak
			Min value	Max value	
[ʦ]	[ʦɿ˥] 資 'capital'	1	3261	9394	7000
		2	3577	9905	7386
		3	3699	10221	6835
	[ʦa˥] 紮 'tie'	1	3212	10416	6118
		2	4064	10416	7607
		3	3090	10489	6780
[ʦʰ]	[ʦʰɿ˥] 疵 'flaw'	1	1995	10489	6780
		2	1703	10586	6174
		3	3139	10513	6174
	[ʦʰa˥] 擦 'rub'	1	3528	9881	6559
		2	4478	9783	6835
		3	4624	10294	6835
[s]	[sɿ˥] 司 'in charge of'	1	3431	10465	6284
		2	3675	10343	6284
		3	2993	10538	7111
	[sa˥] 撒 'cast'	1	2896	9856	4465
		2	5889	10489	8654
		3	2506	10319	4354
[tʂ]	[tʂʅ˥] 知 'know'	1	1971	10489	3362
		2	2677	10440	4079
		3	1119	10270	3858
	[tʂa˥] 渣 'residue'	1	5816	9686	7276
		2	5670	9735	6118
		3	6011	10513	7111
[tʂʰ]	[tʂʰʅ˥] 吃 'eat'	1	2360	10392	2921
		2	2093	10562	3969
		3	2482	10586	3472
	[tʂʰa˥] 叉 'folk'	1	3139	10416	4354
		2	2969	10416	4299
		3	2896	9929	4299

Target sibilant	Test words	Token no.	Noise range		Noise peak
			Min value	Max value	
[ʂ]	[ʂʅ˥] 失 'lose'	1	2798	10489	3969
		2	2287	10465	3858
		3	2239	10562	3913
	[ʂa˥] 沙 'sand'	1	3650	9978	7221
		2	3918	9905	7386
		3	4210	10294	7276
[tɕ]	[tɕi˥] 基 'base'	1	6790	9394	8379
		2	5670	9004	8820
		3	5622	10100	6835
	[tɕia˥] 加 'add'	1	3188	9248	6780
		2	2847	10489	7111
		3	2482	9856	6229
[tɕʰ]	[tɕʰi˥] 七 'seven'	1	2774	10538	5512
		2	3382	9321	6559
		3	2385	10513	6284
	[tɕʰia˥] 掐 'nip off'	1	2190	9126	6945
		2	3188	10440	6284
		3	2823	10173	5016
[ɕ]	[ɕi˥] 希 'hope'	1	2360	9370	6339
		2	2263	9613	6890
		3	2628	10465	5953
	[ɕia˥] 蝦 'prawn'	1	2628	10075	6174
		2	4405	10027	5953
		3	2336	10538	4630

Appendix 3: Wide-band Spectrograms and LPC&FFT Spectra

3.1 Native speaker

Token 1

Token 2

Token 3

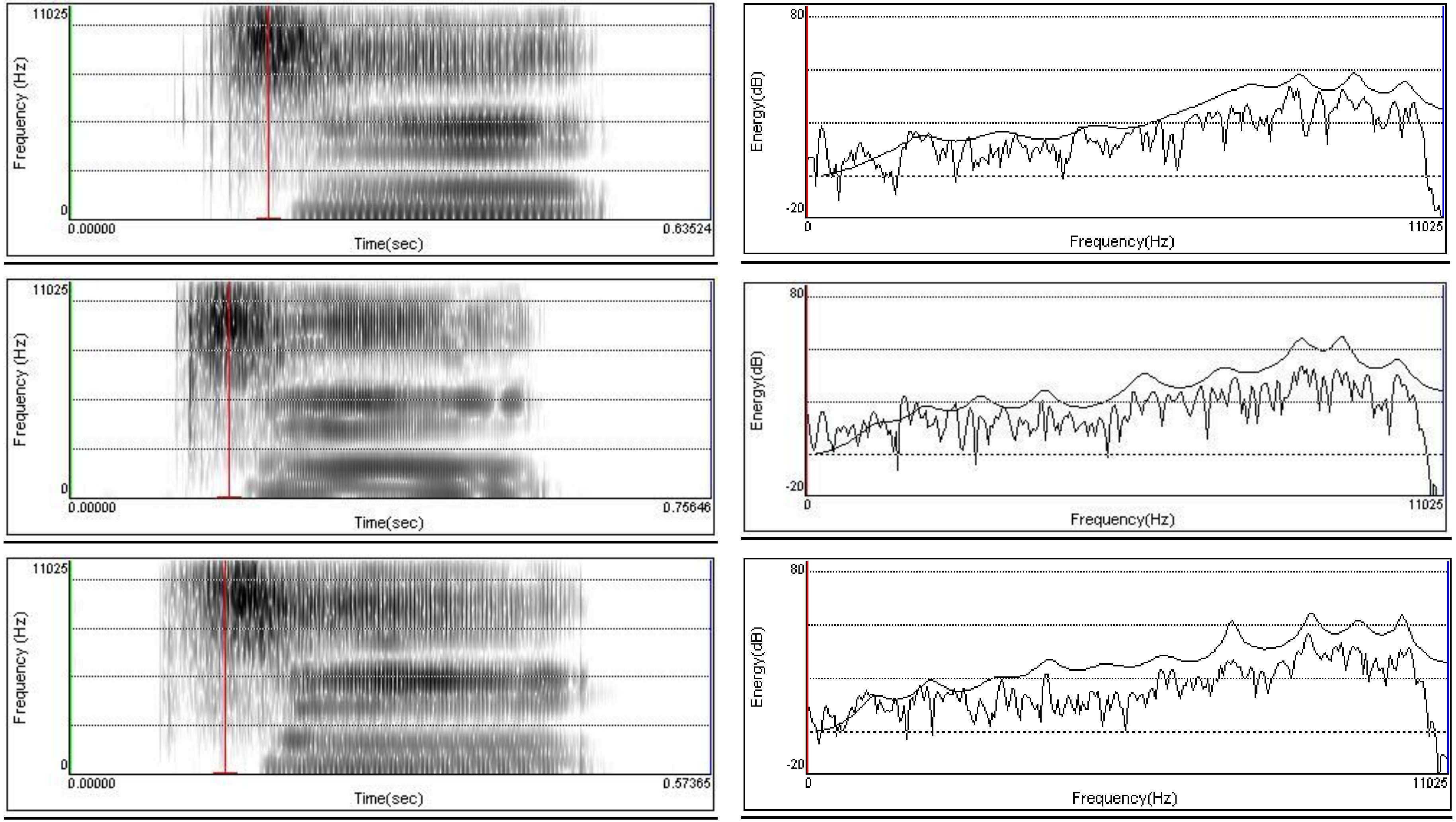


Figure 3.1.1. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [ʦ] in [ʦɿ˥] 資 ‘capital’

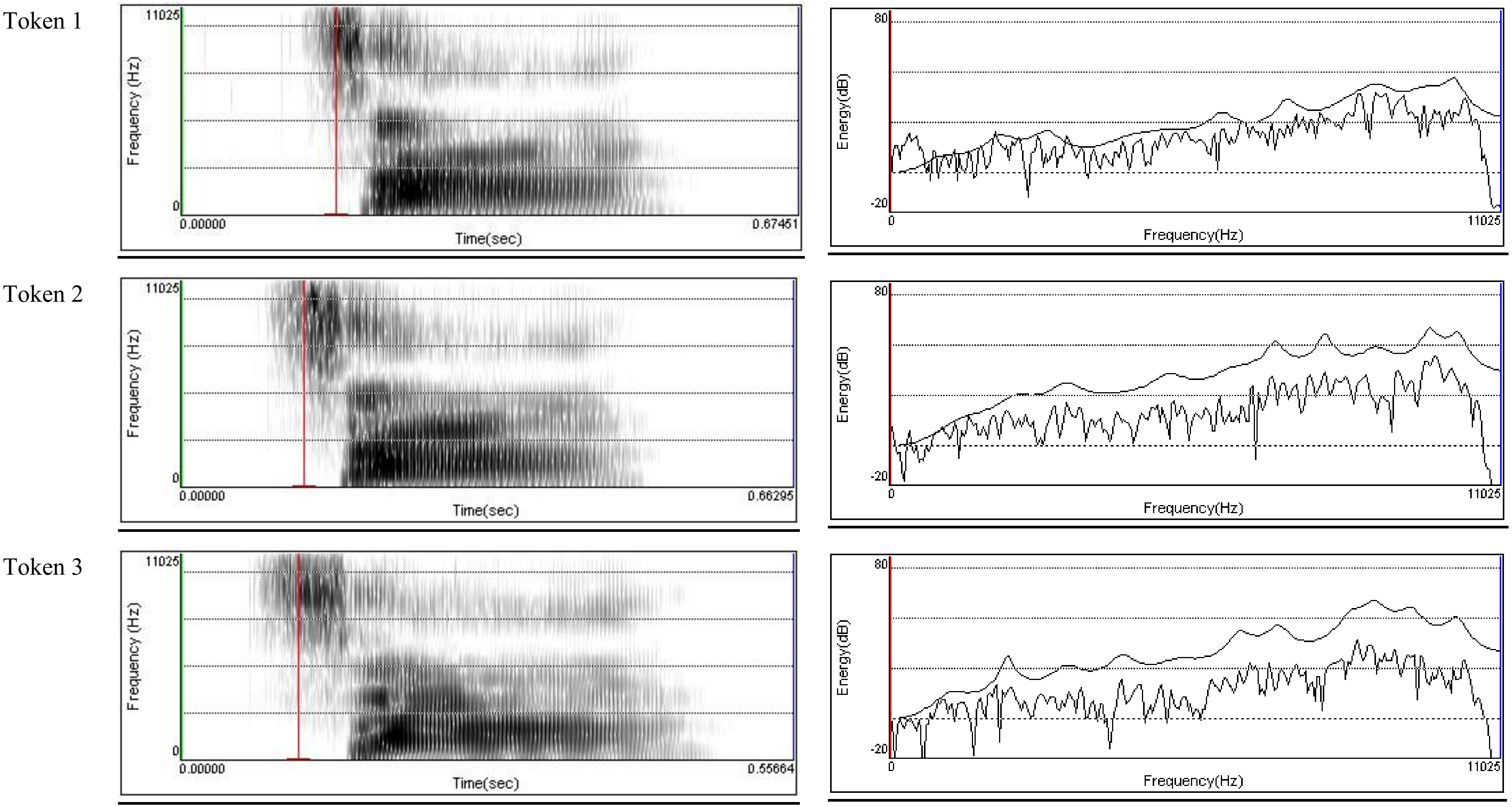


Figure 3.1.2. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [ʦ] in [ʦa˥] 紮 ‘tie’

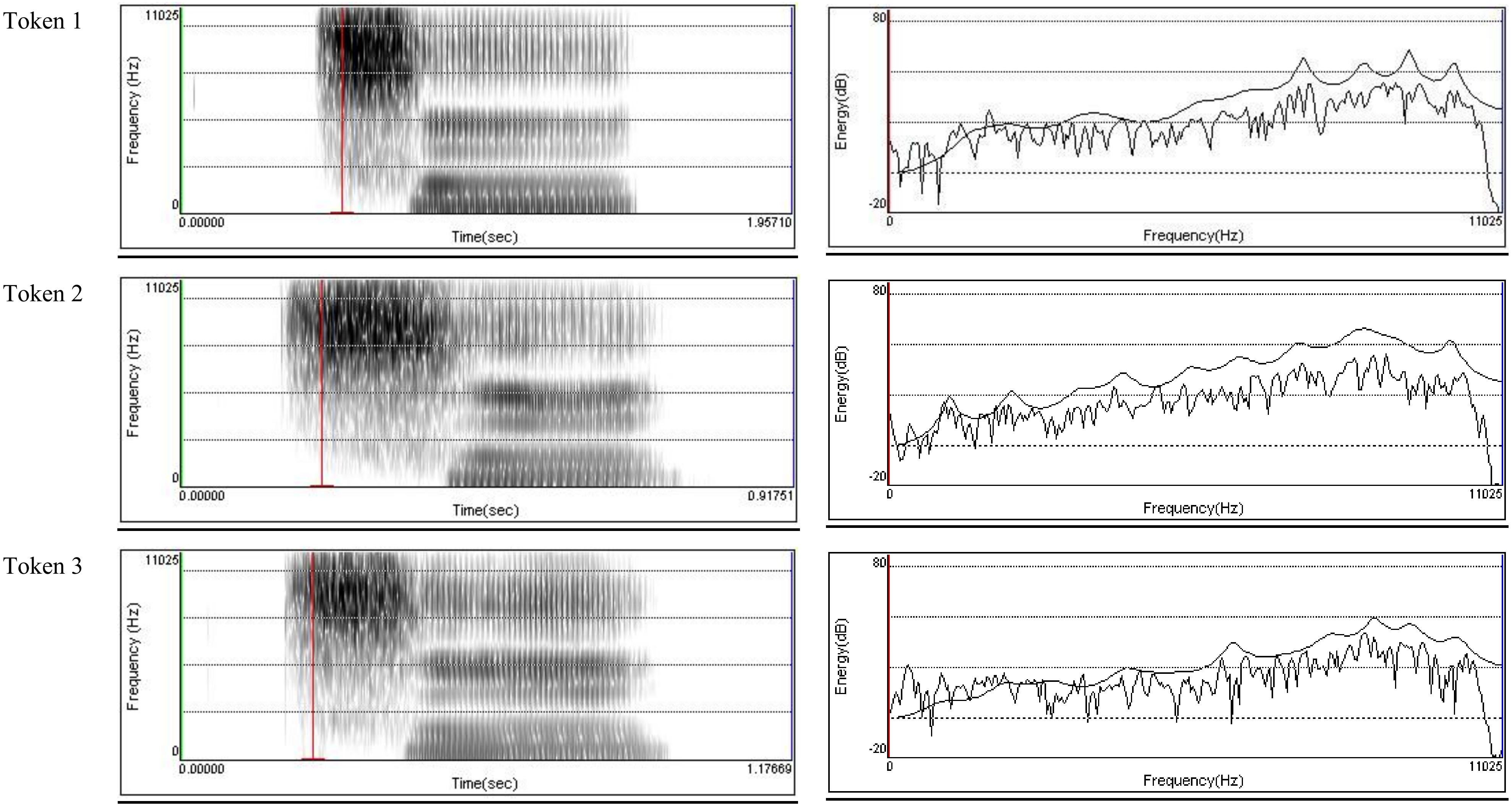


Figure 3.1.3. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [ts^h] in [ts^{h}ɿ˥] 疵 ‘flaw’

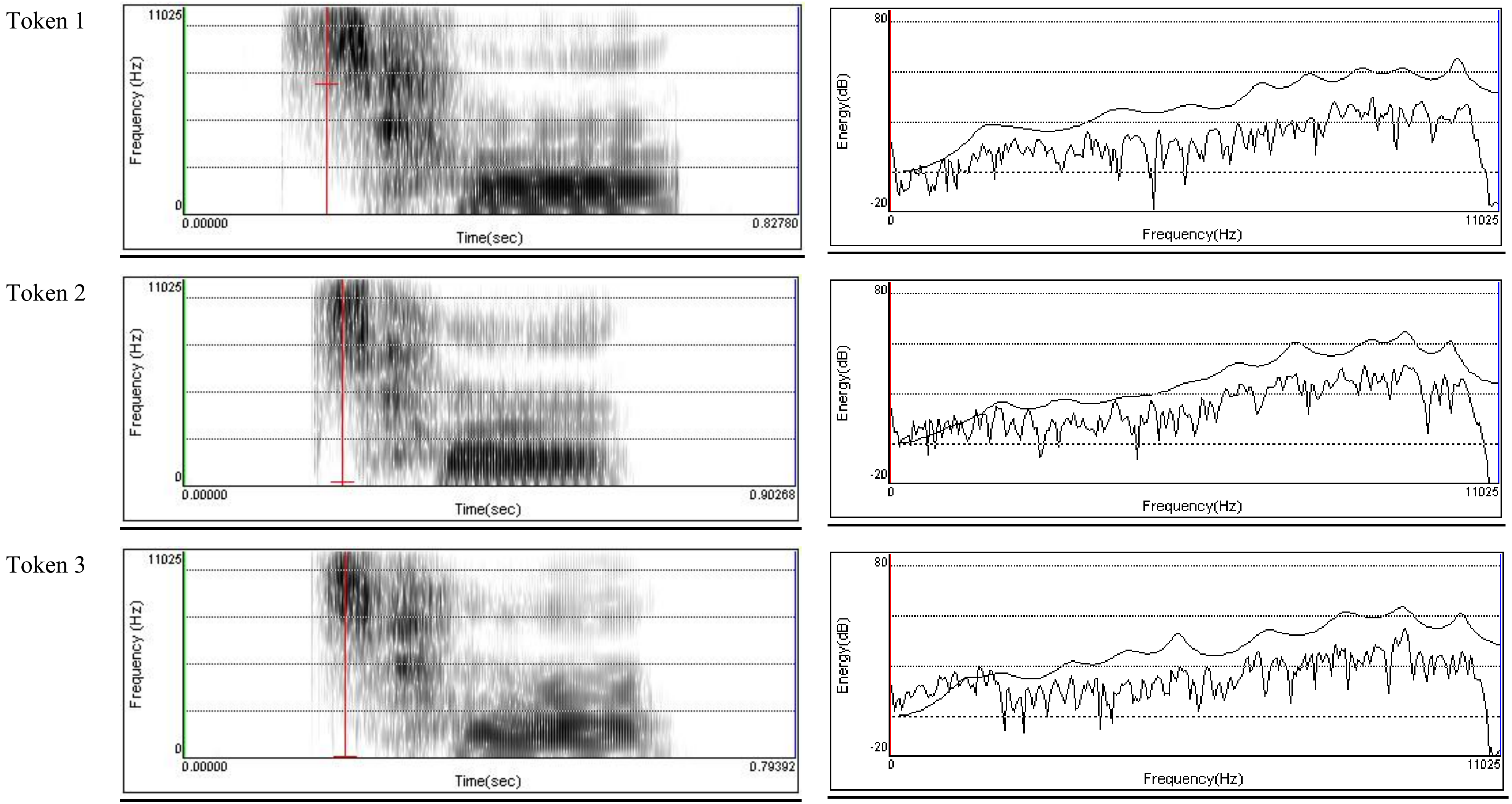


Figure 3.1.4. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [ʦ^{h}] in [ʦ^{h}a˥] 擦 ‘rub’

Token 1

Token 2

Token 3

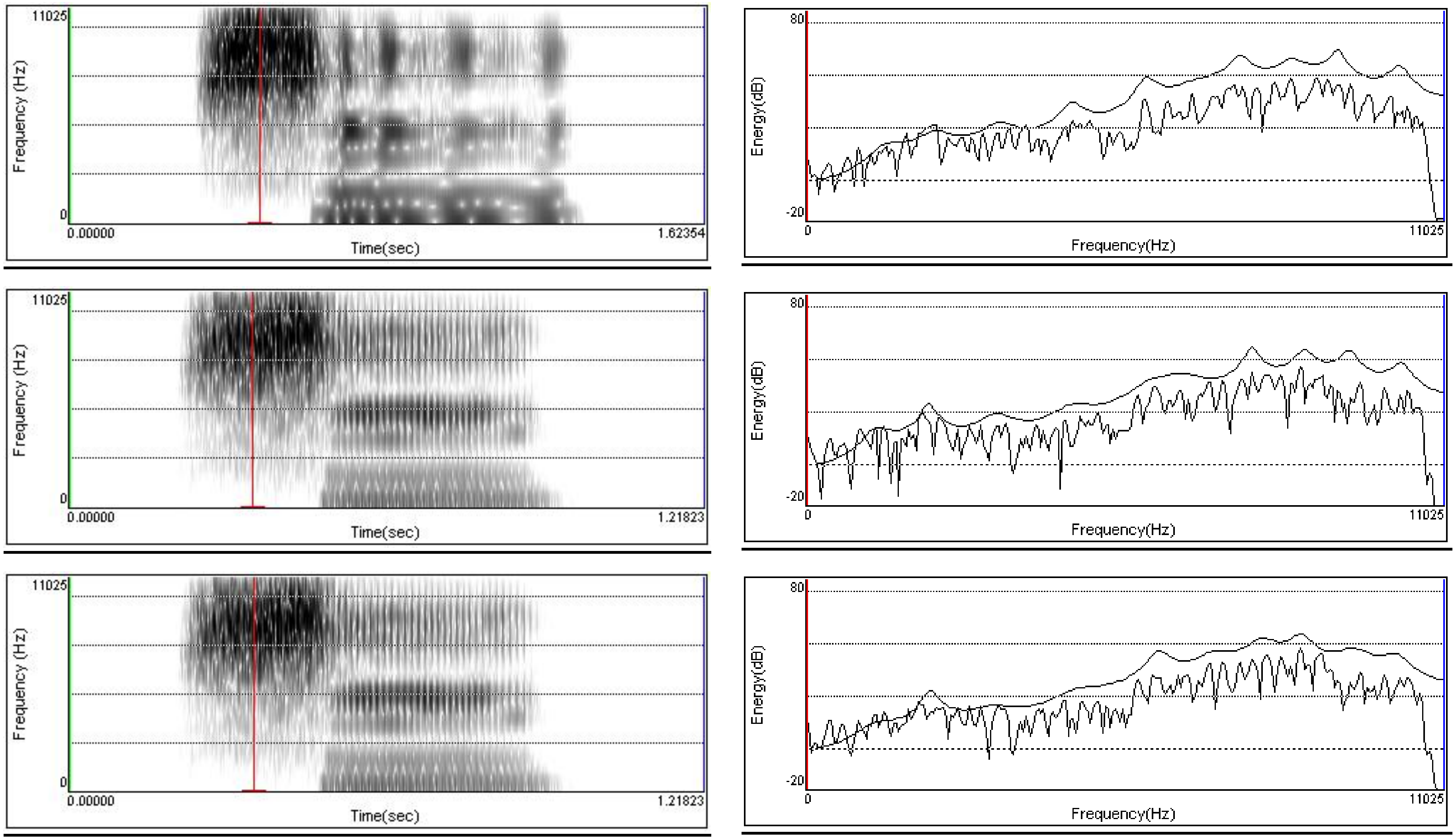


Figure 3.1.5. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [s] in [sɿ˥] 司 ‘in charge of’

Token 1

Token 2

Token 3

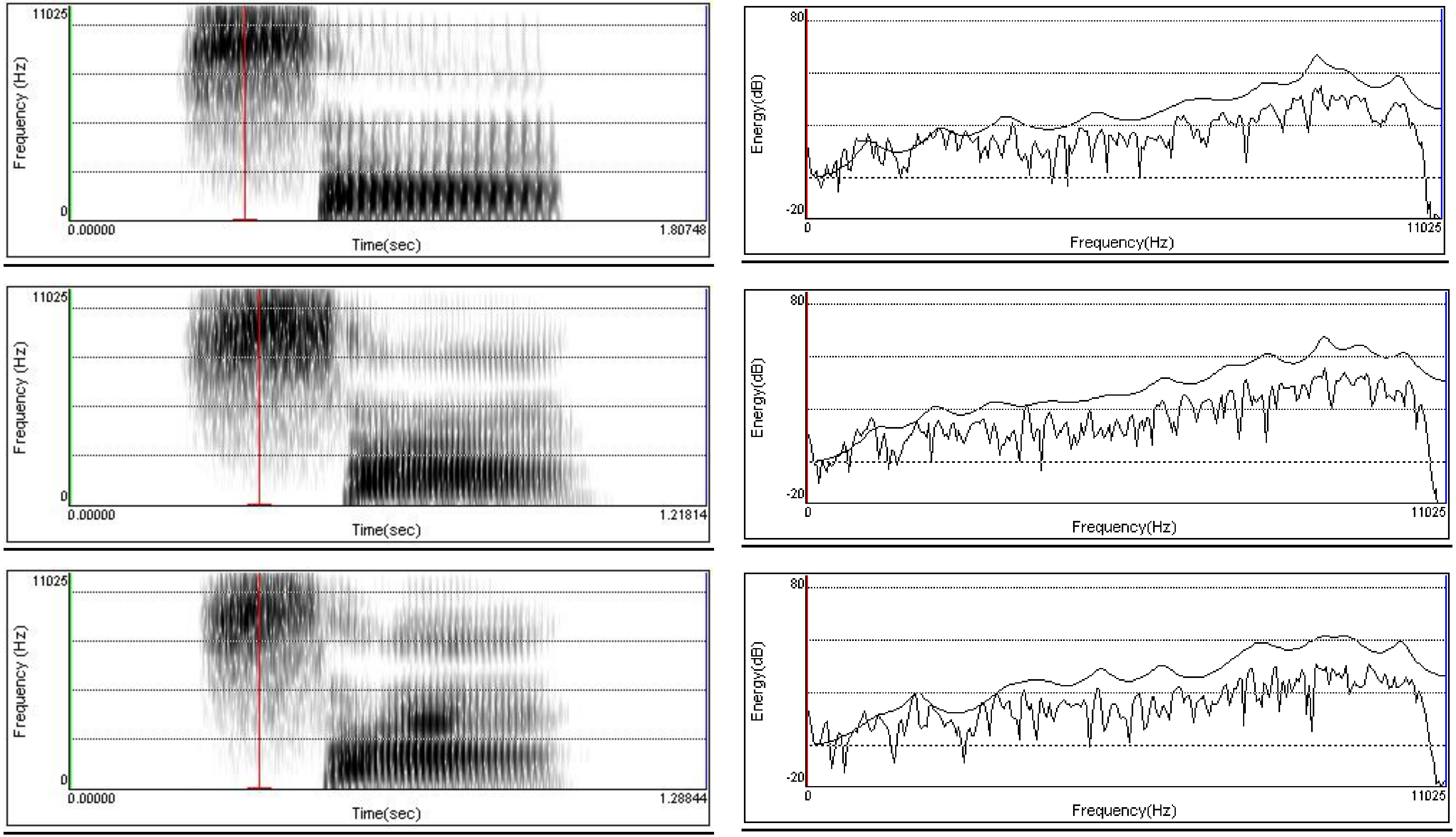


Figure 3.1.6. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [s] in [sa˥] 撒 ‘cast’

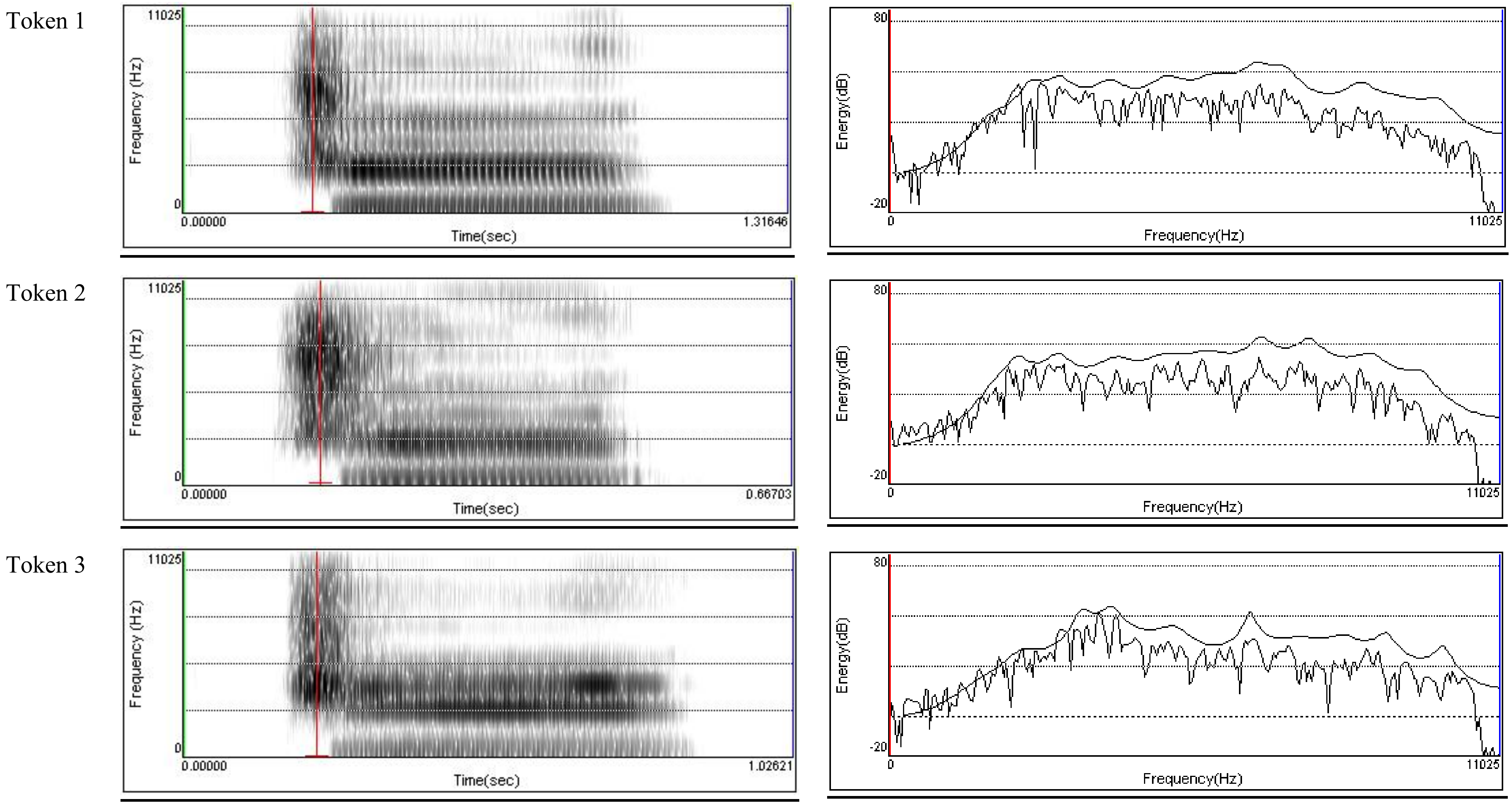


Figure 3.1.7. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [tʂ] in [tʂʅ˥] 知 'know'

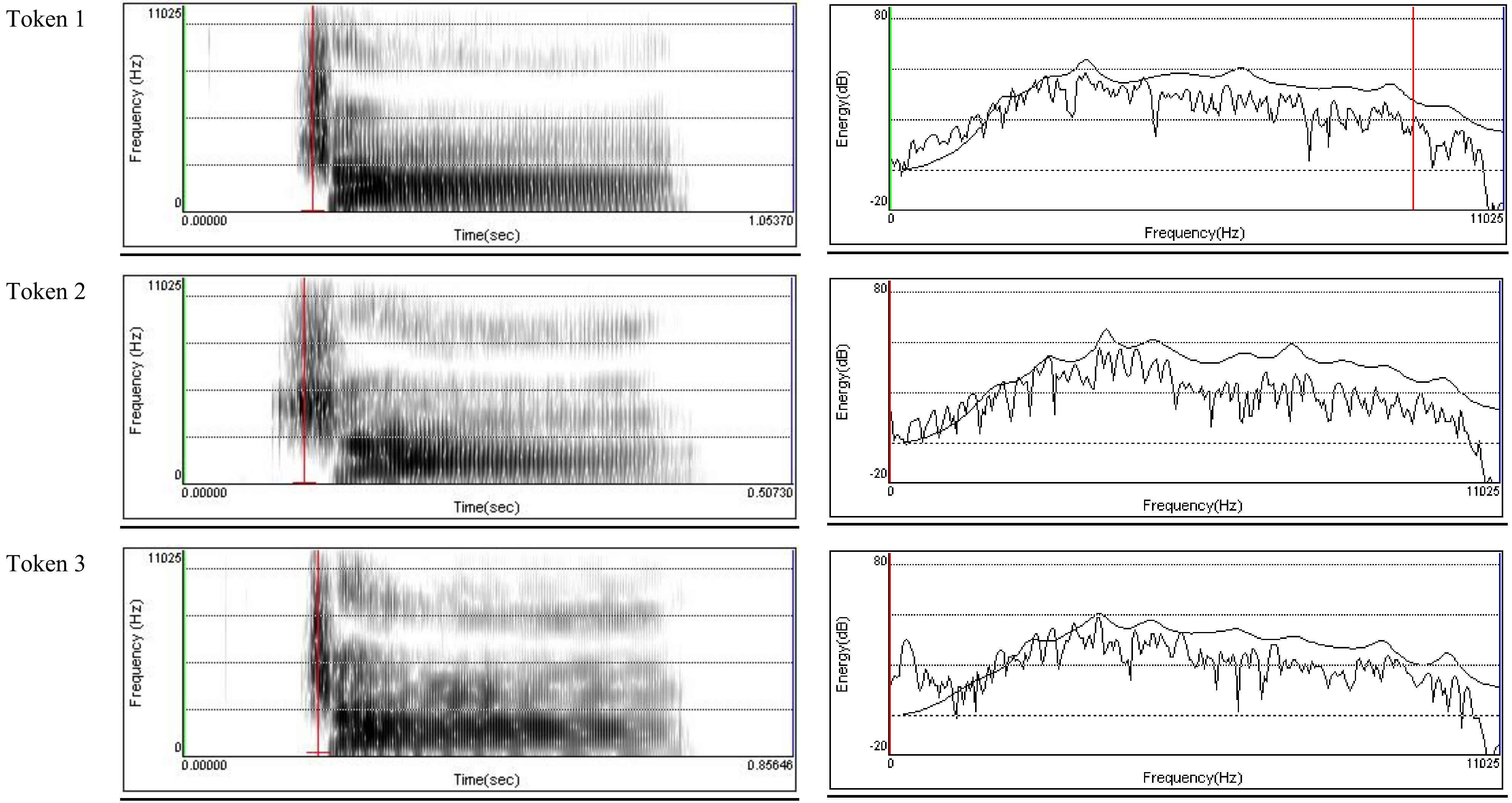


Figure 3.1.8. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [tʂ] in [tʂa˥] 渣 ‘residue’

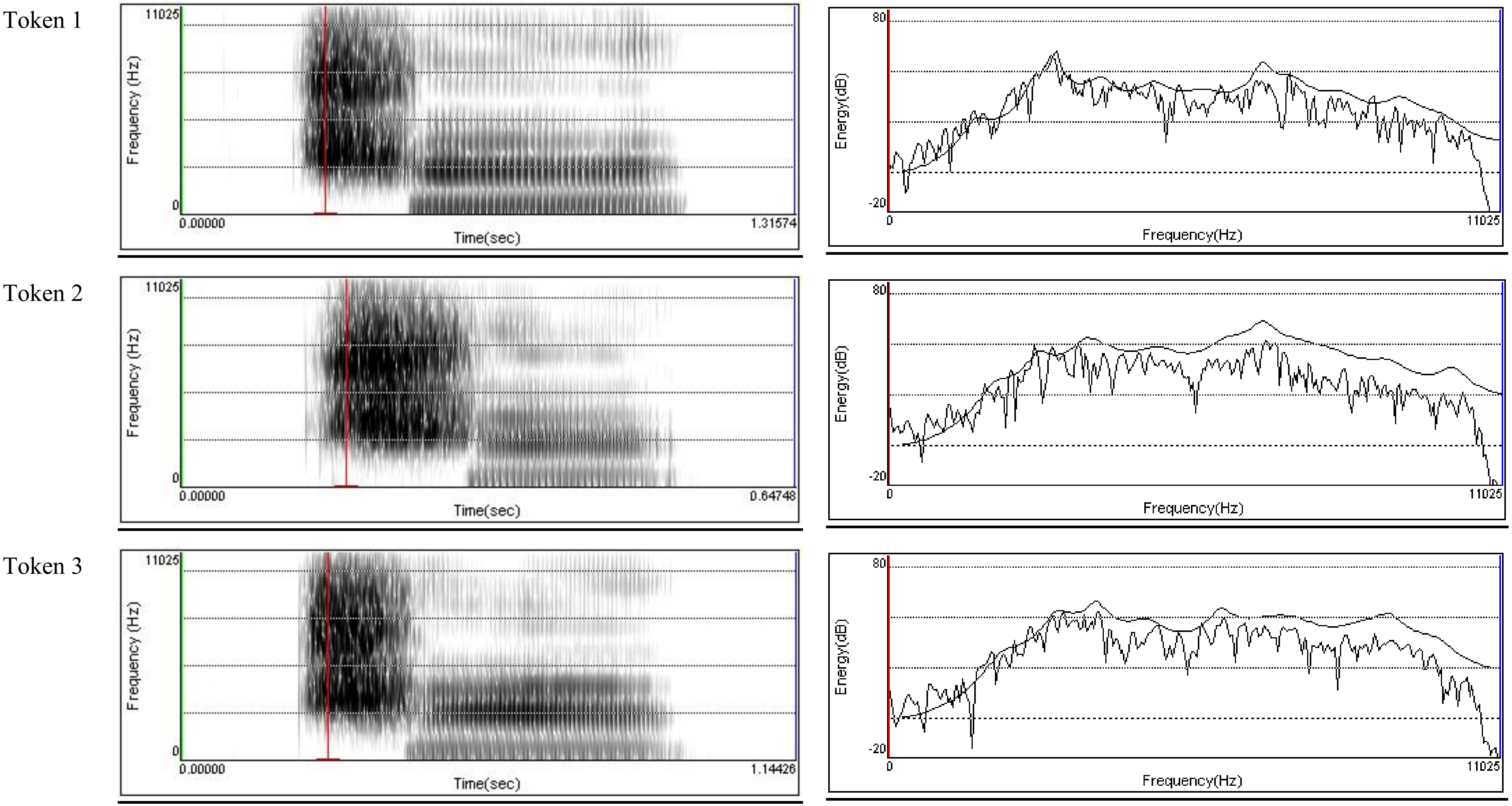


Figure 3.1.9. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [tʂʰ] in [tʂʰʅ˥] 吃 ‘eat’

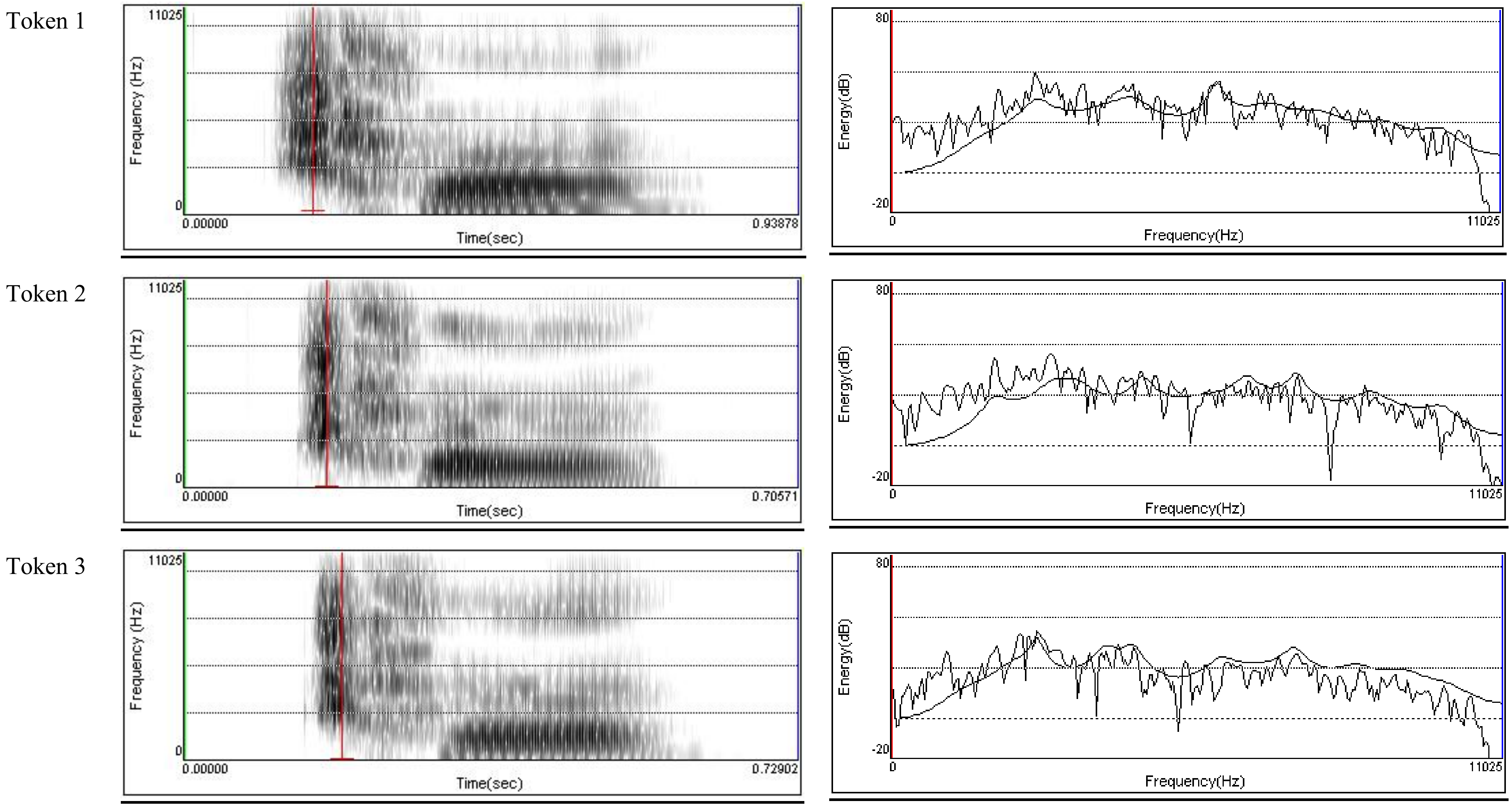


Figure 3.1.10. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [tʂʰ] in [tʂʰa˥] 叉 ‘folk’

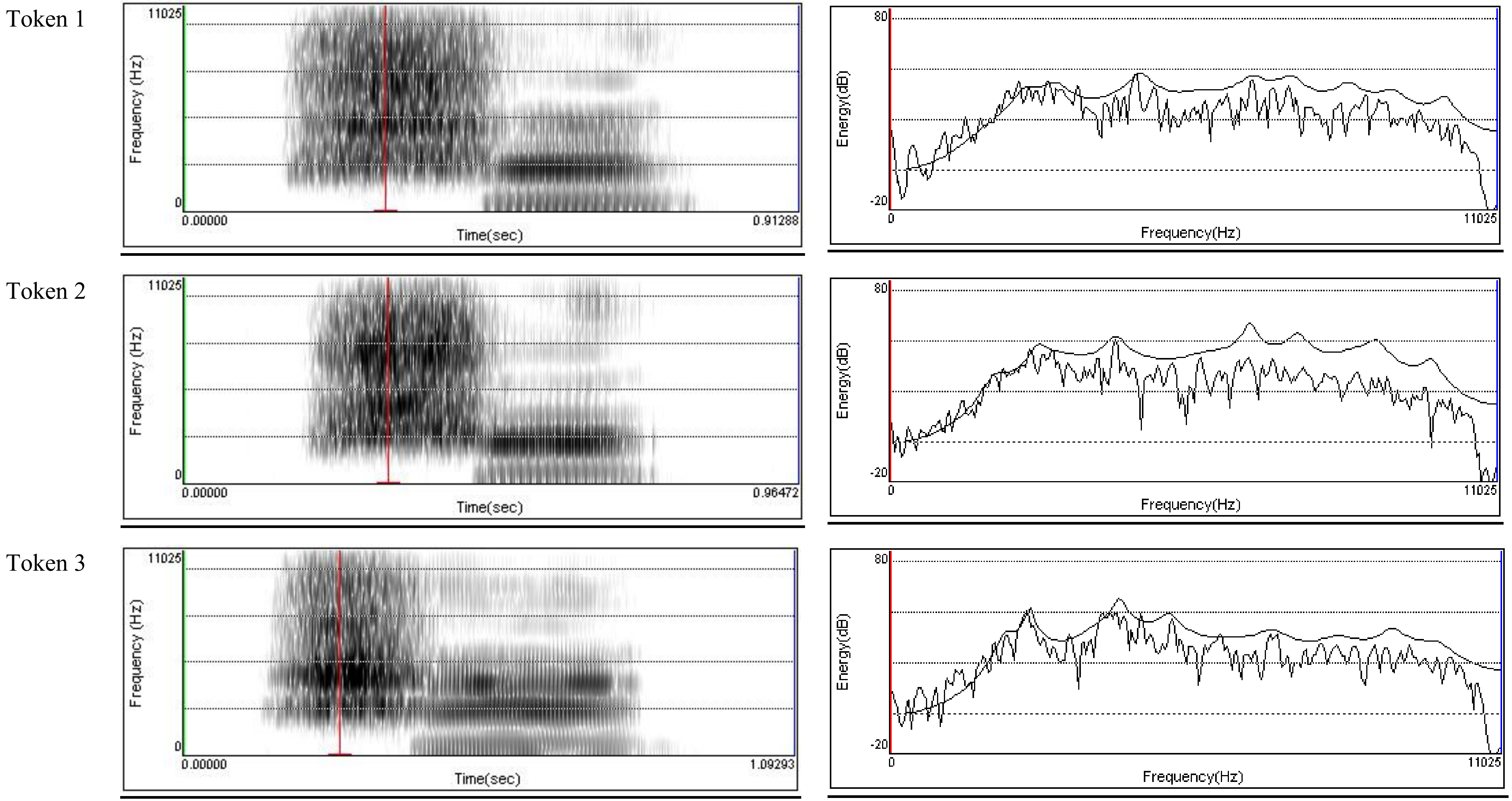


Figure 3.1.11. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [ʂ] in [ʂʅ˥] 失 'lose'

Token 1

Token 2

Token 3

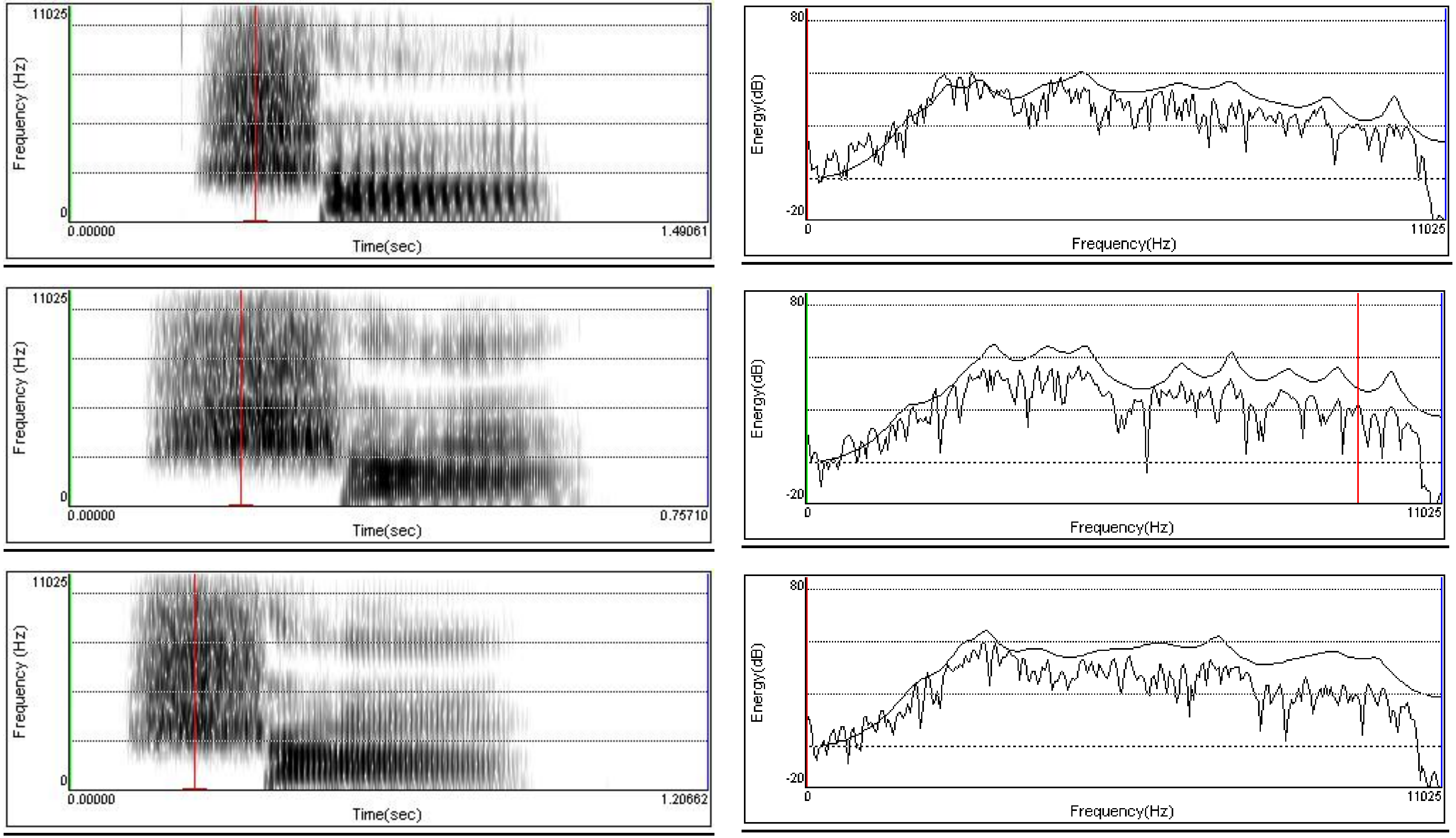


Figure 3.1.12. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [ʂ] in [ʂa˥] 沙 'sand'

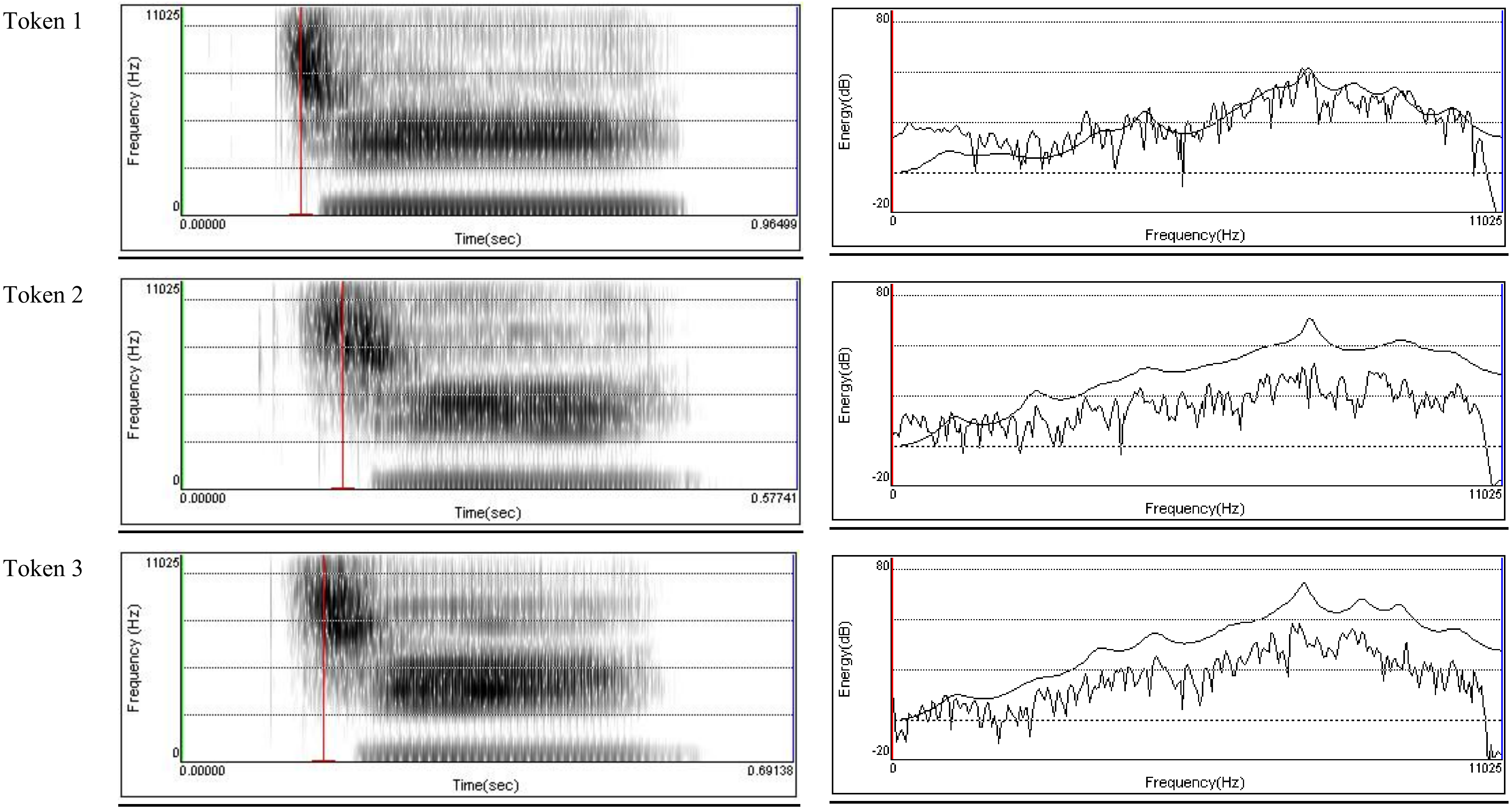


Figure 3.1.13. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [tɕ] in [tɕi˥] 基 ‘base’

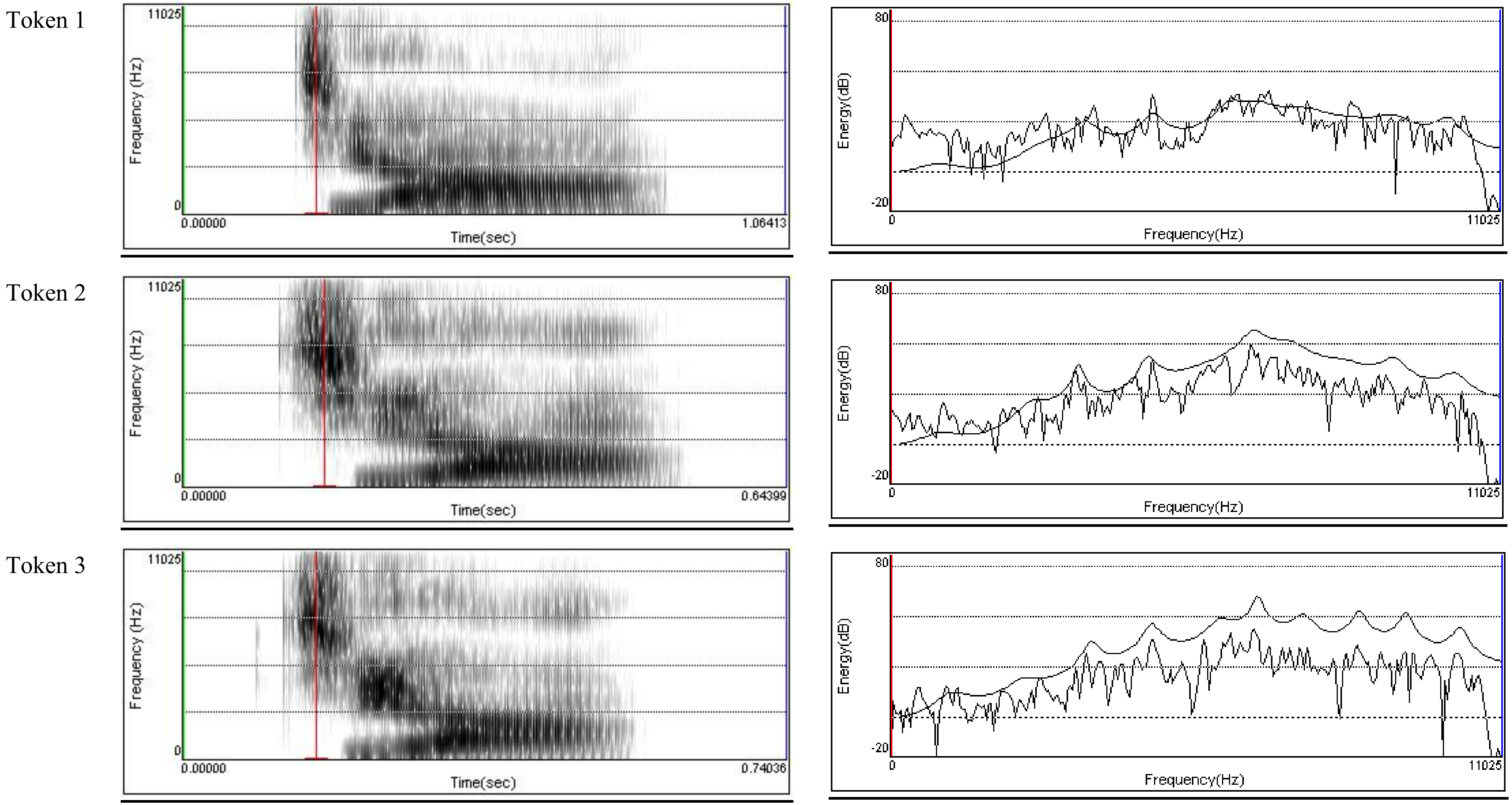


Figure 3.1.14. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [tɕ] in [tɕia˥] 加 'add'

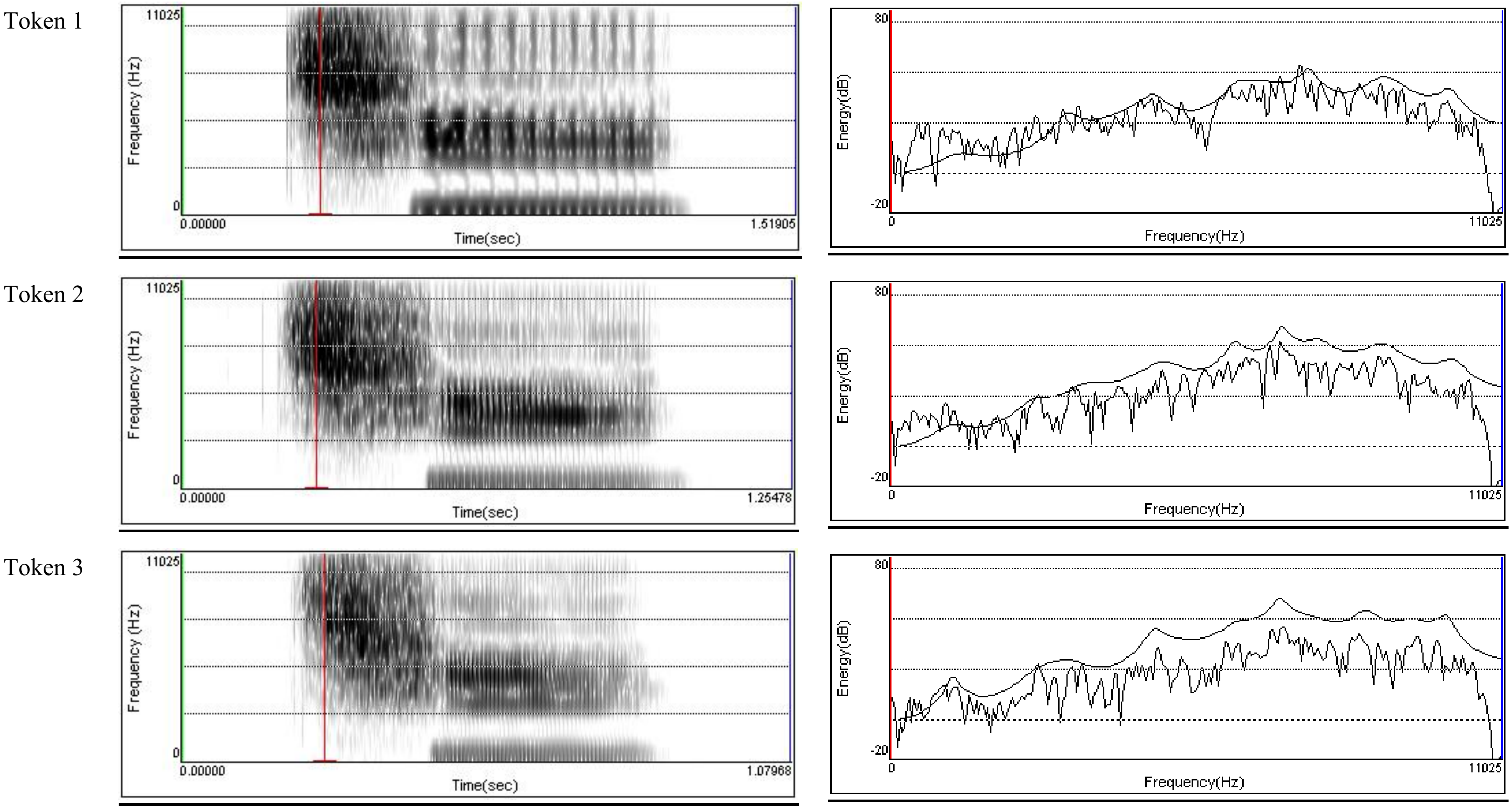


Figure 3.1.15. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [tɕʰ] in [tɕʰi˥] 七 'seven'

Token 1

Token 2

Token 3

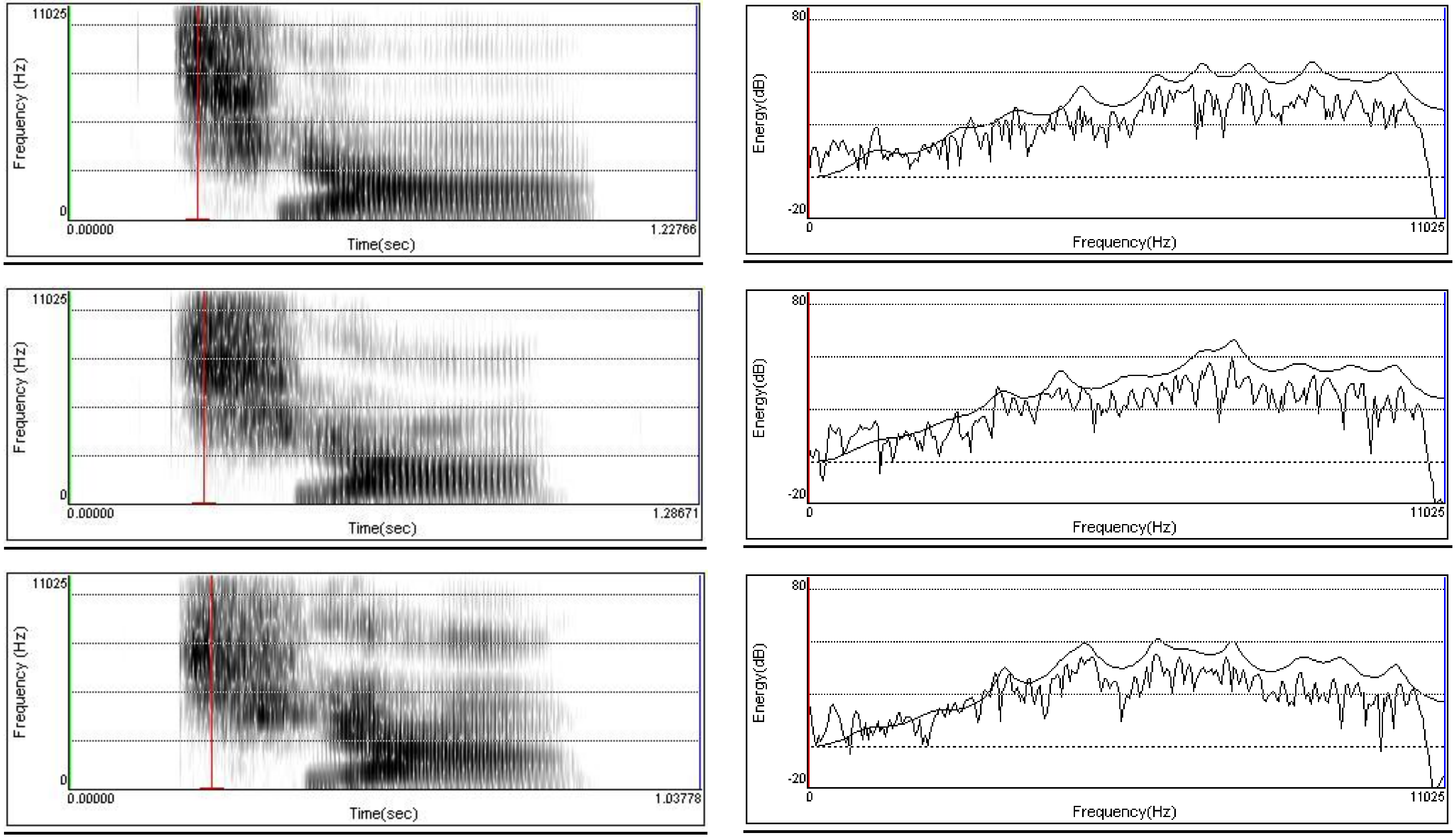


Figure 3.1.16. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [tɕʰ] in [tɕʰia˥] 掐 ‘nip off’

Token 1

Token 2

Token 3

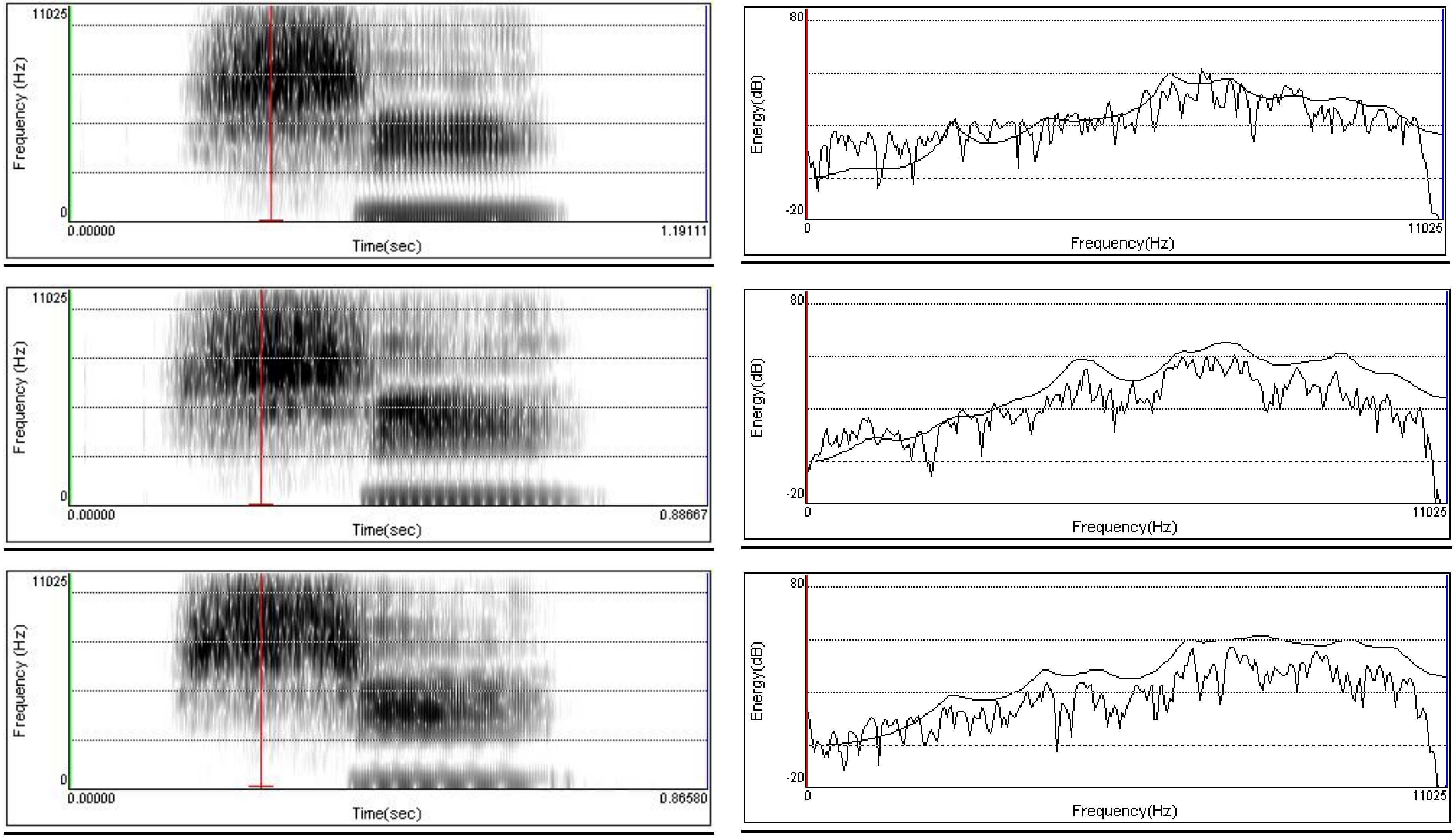


Figure 3.1.17. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [ɕ] in [ɕi˥] 希 ‘hope’

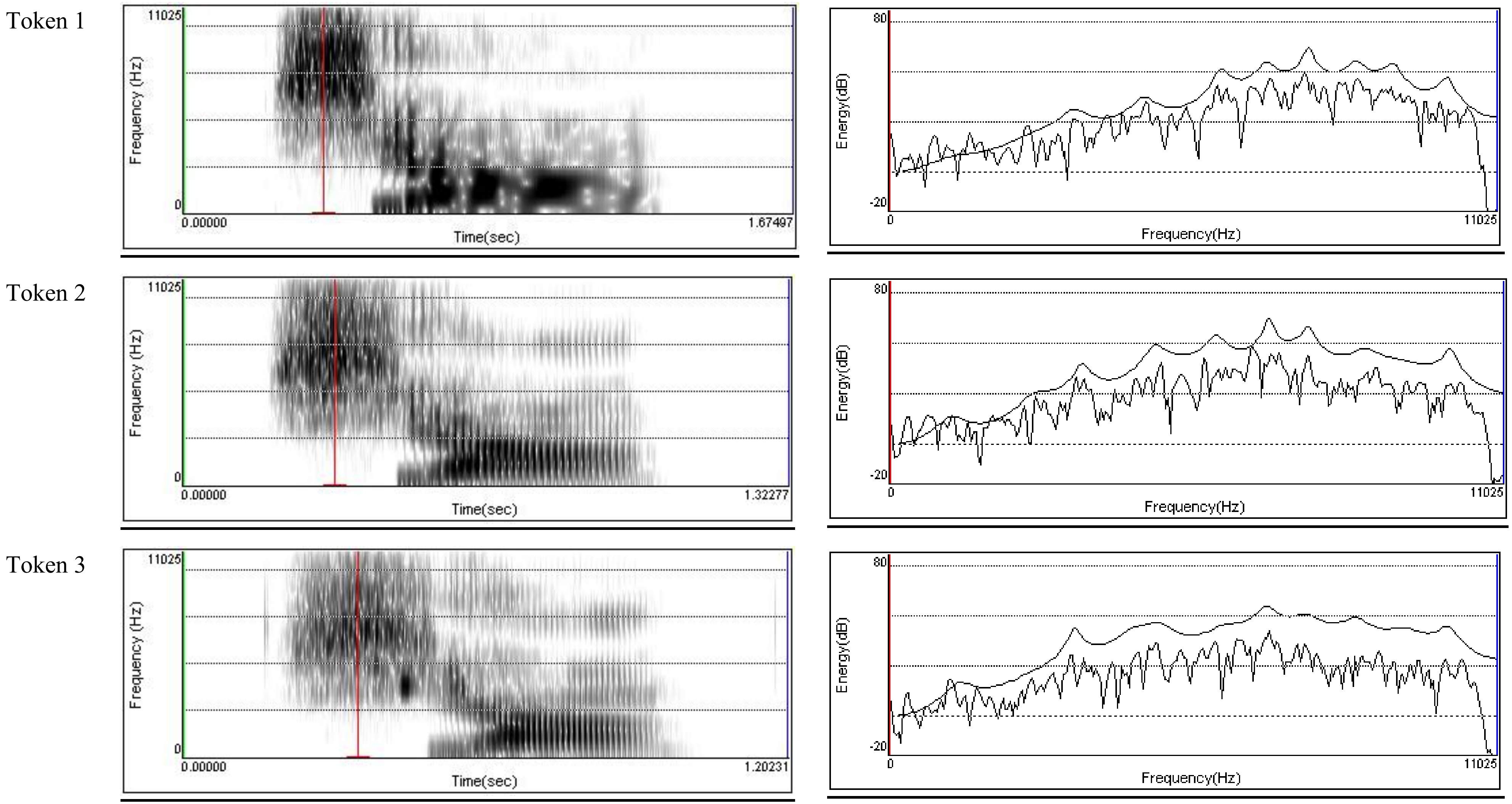


Figure 3.1.18. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [ɕ] in [ɕia˥] 蝦 'prawn'

3.2 Cantonese Male 1

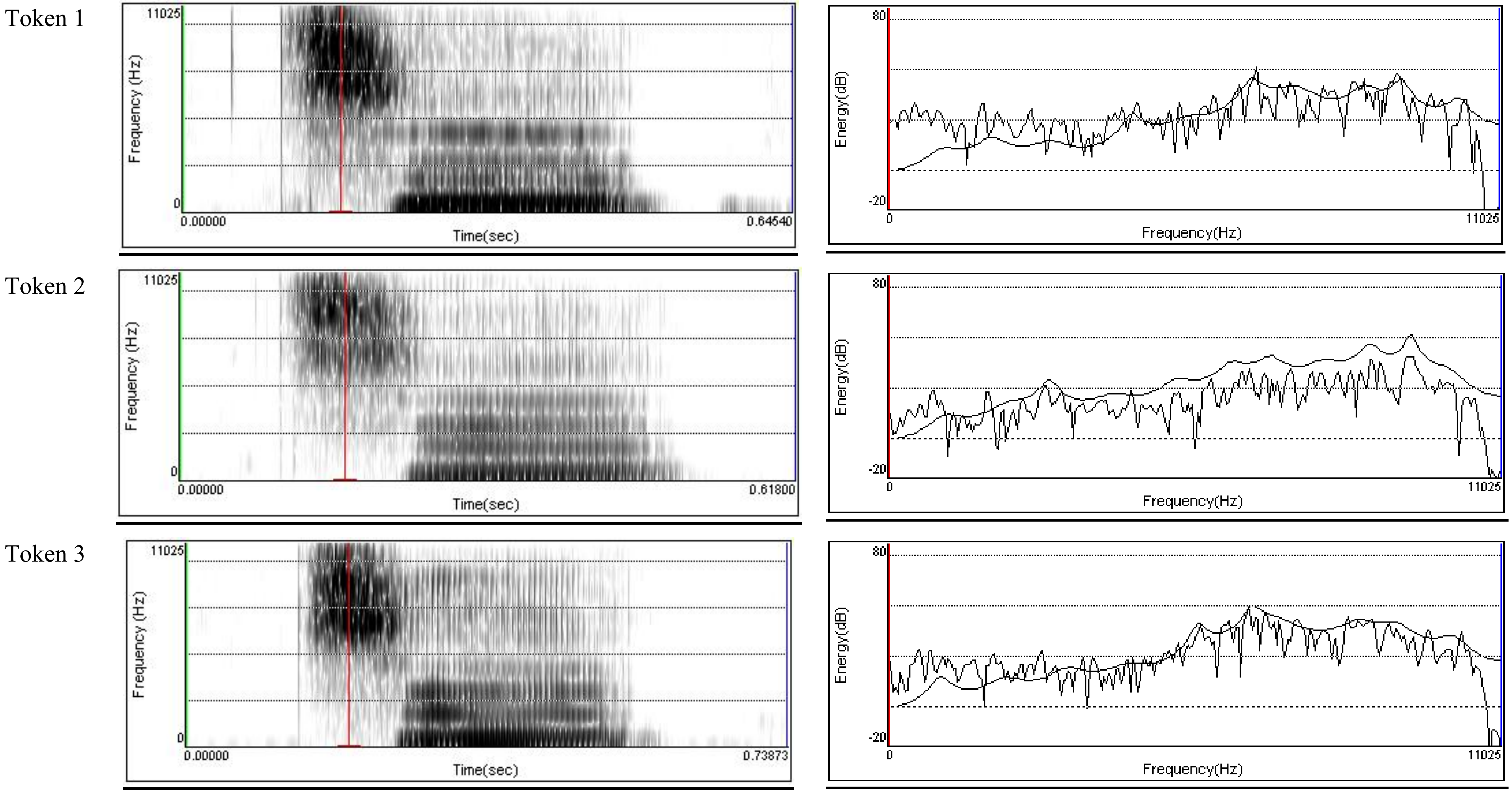


Figure 3.2.1. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [ʦ] in [ʦɿ˥] 資 ‘capital’

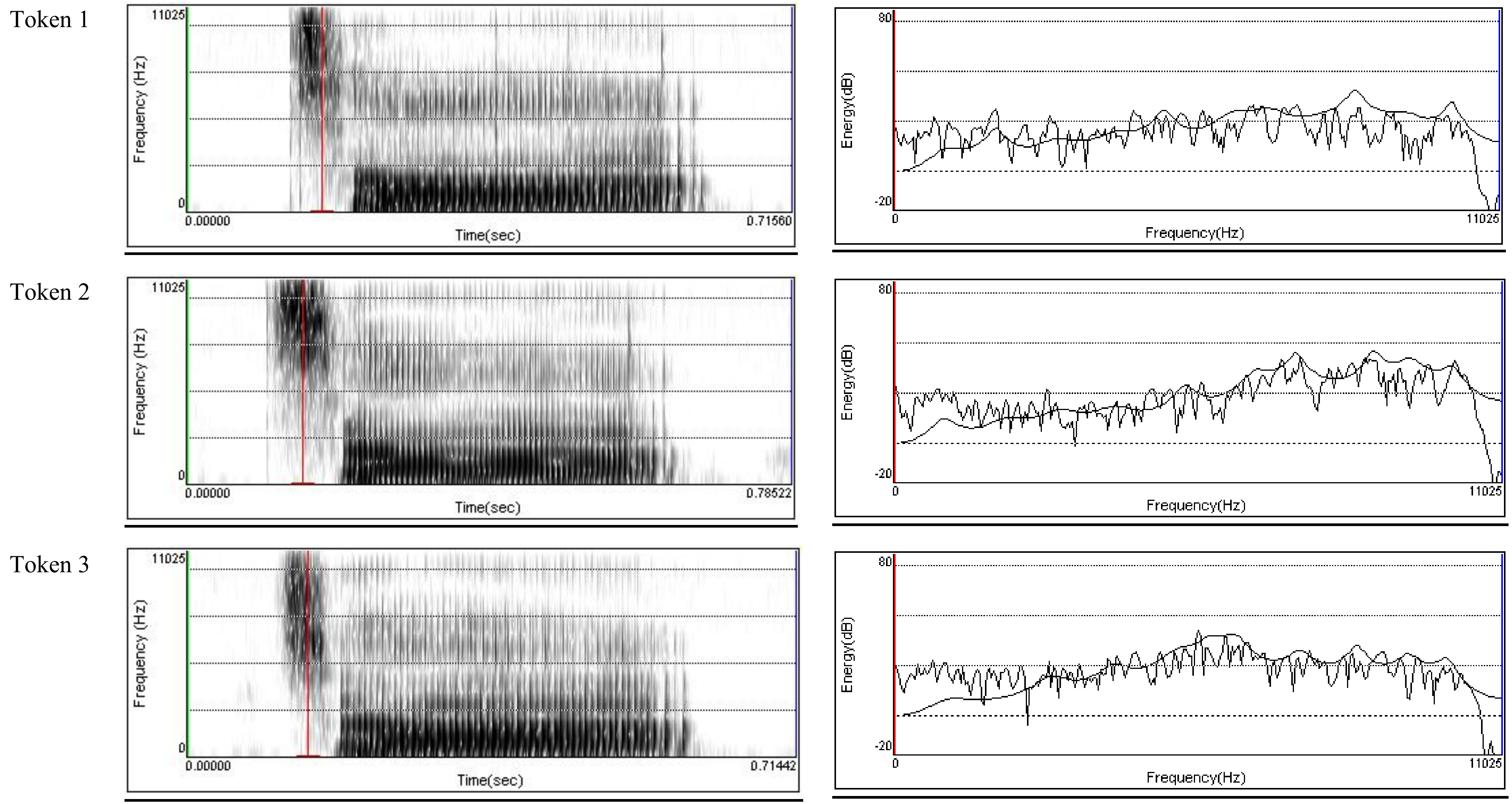


Figure 3.2.2. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [ʦ] in [ʦa˥] 紮 'tie'

Token 1

Token 2

Token 3

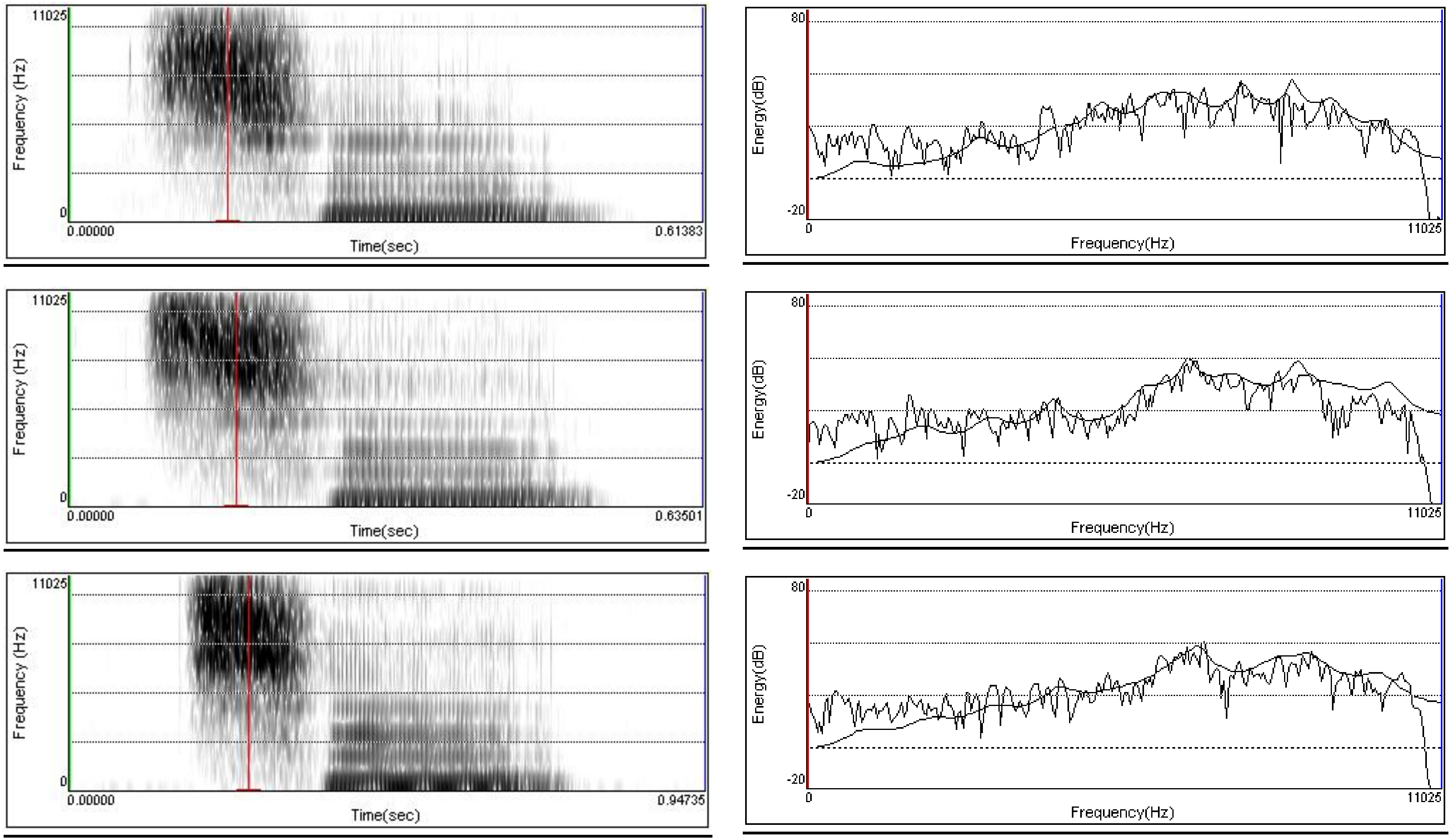


Figure 3.2.3. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [ʦ^h] in [ʦ^{h}ɿ˥] 疵 ‘flaw’

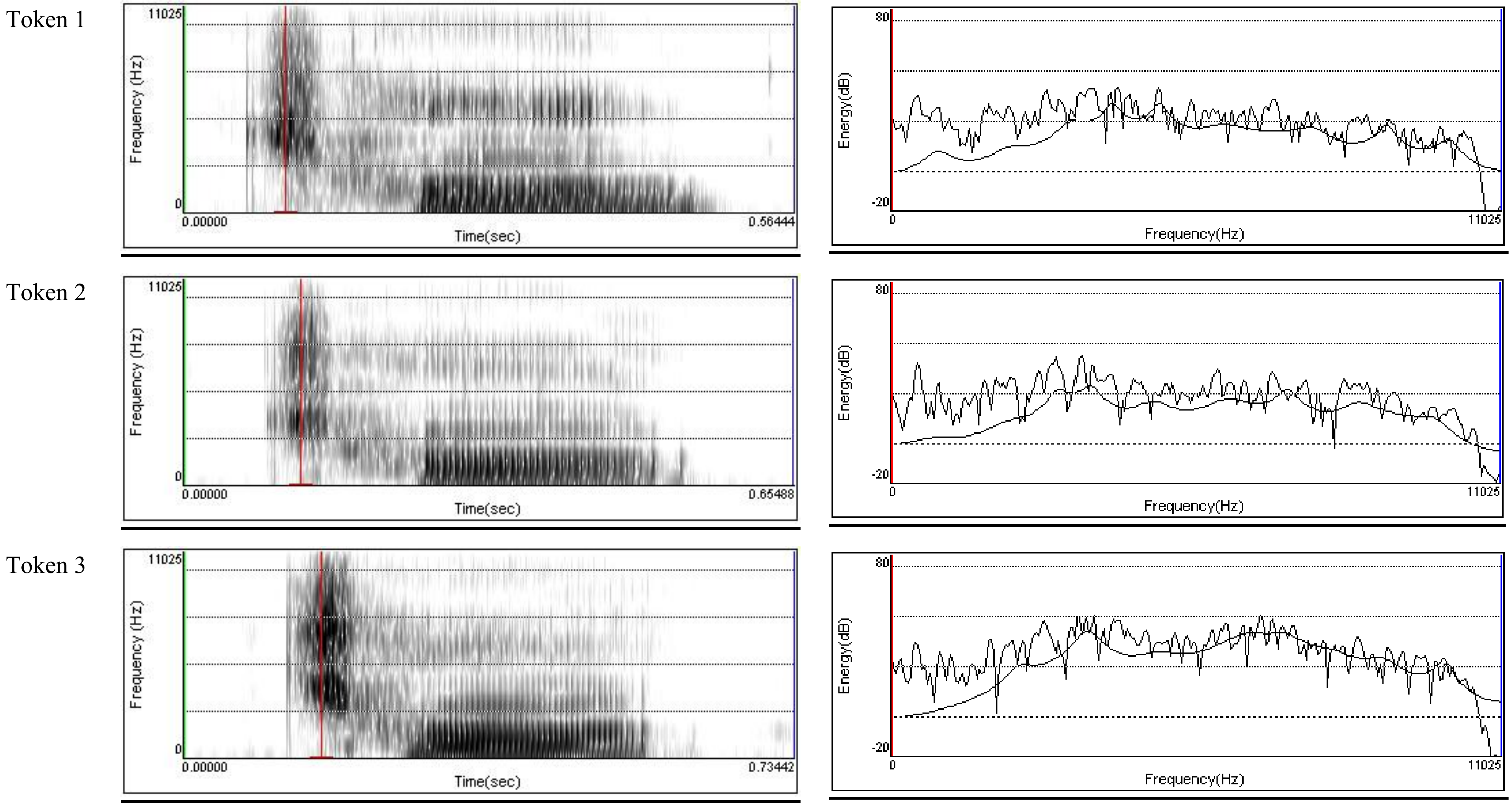


Figure 3.2.4. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [tsʰ] in [tsʰa˥] 擦 ‘rub’

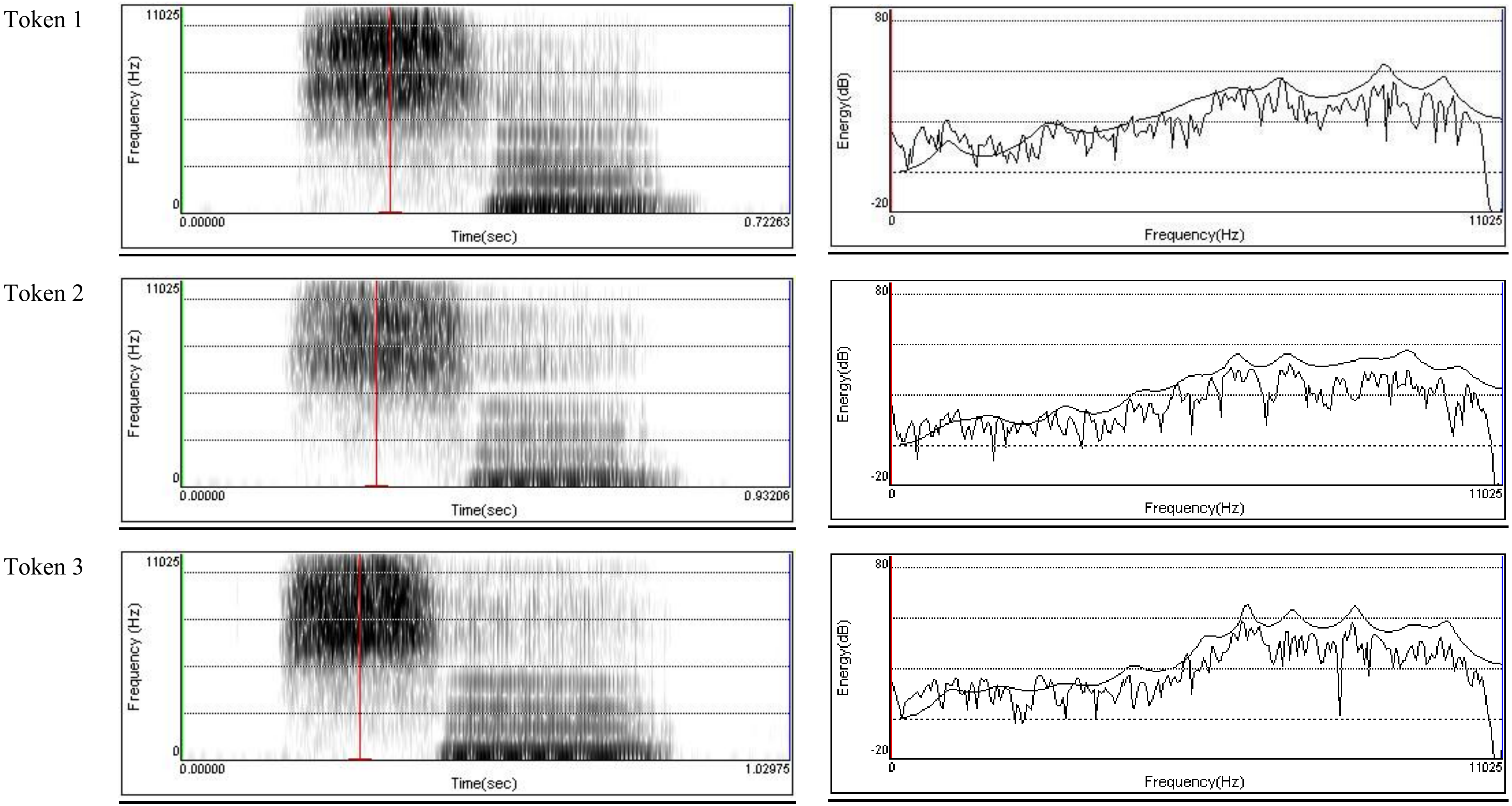


Figure 3.2.5. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [s] in [sɿ˥] 司 ‘in charge of’

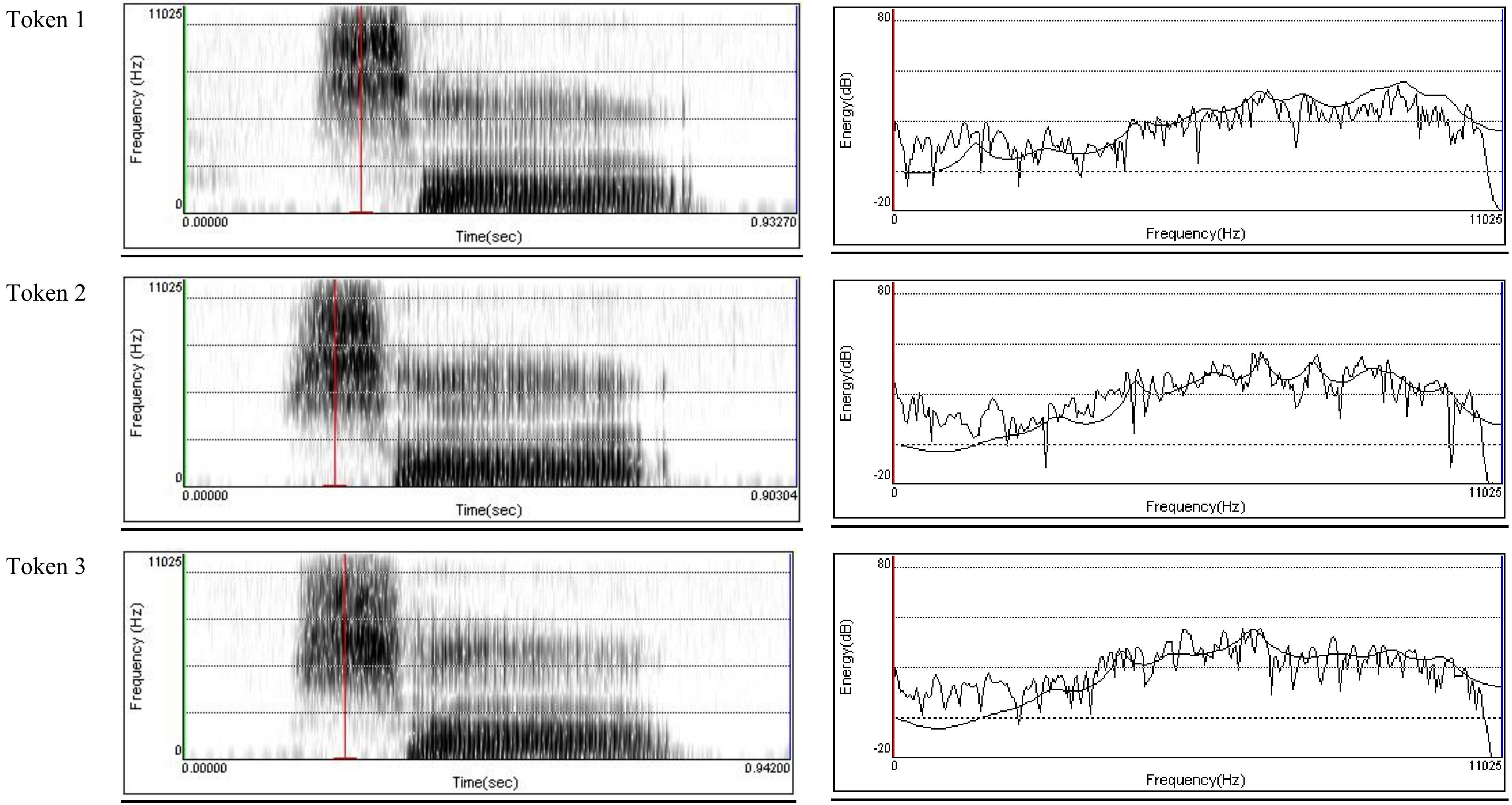


Figure 3.2.6. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [s] in [sa˥] 撒 ‘cast’

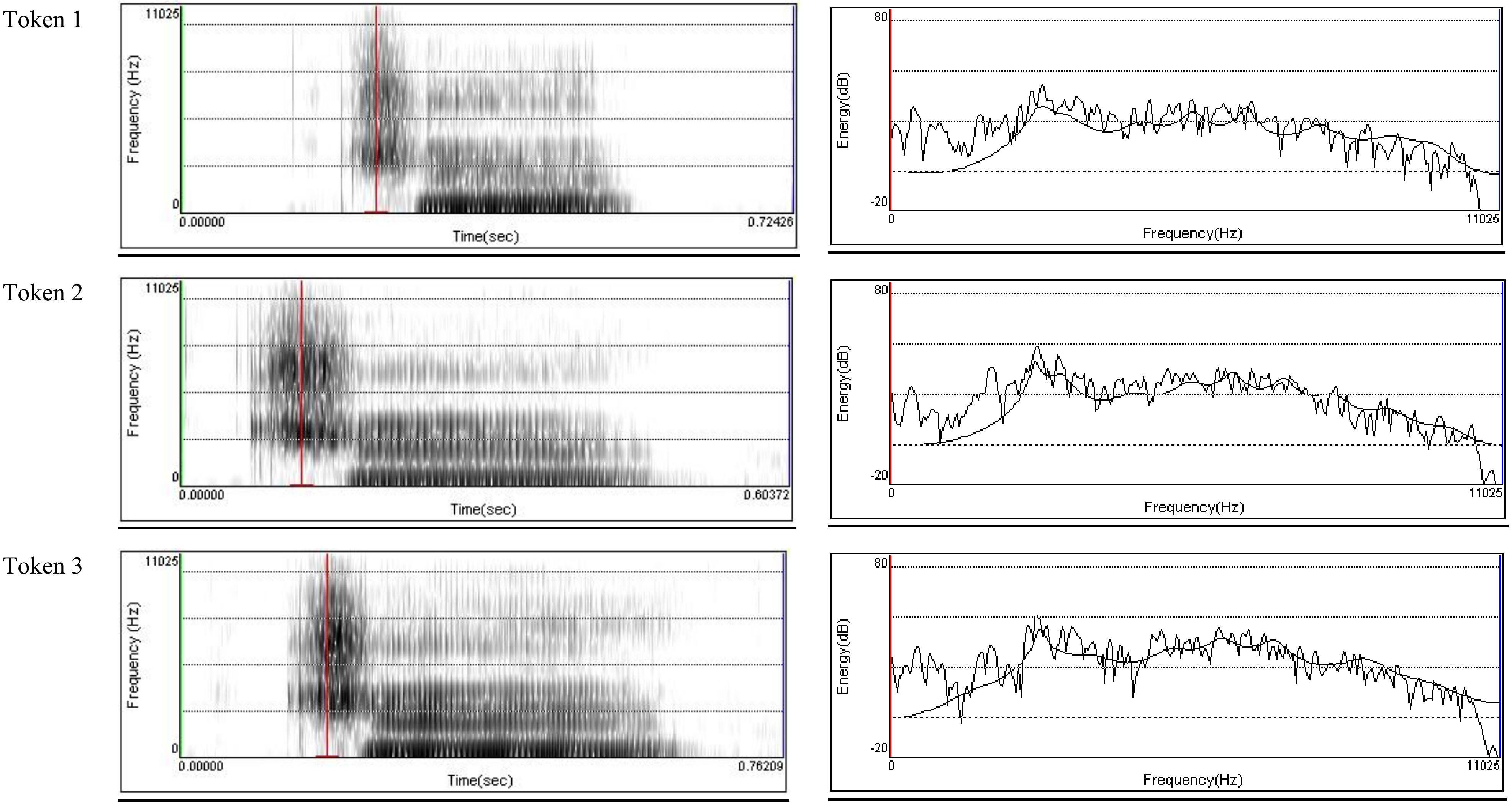


Figure 3.2.7. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [tʂ] in [tʂʅ˥] 知 ‘know’

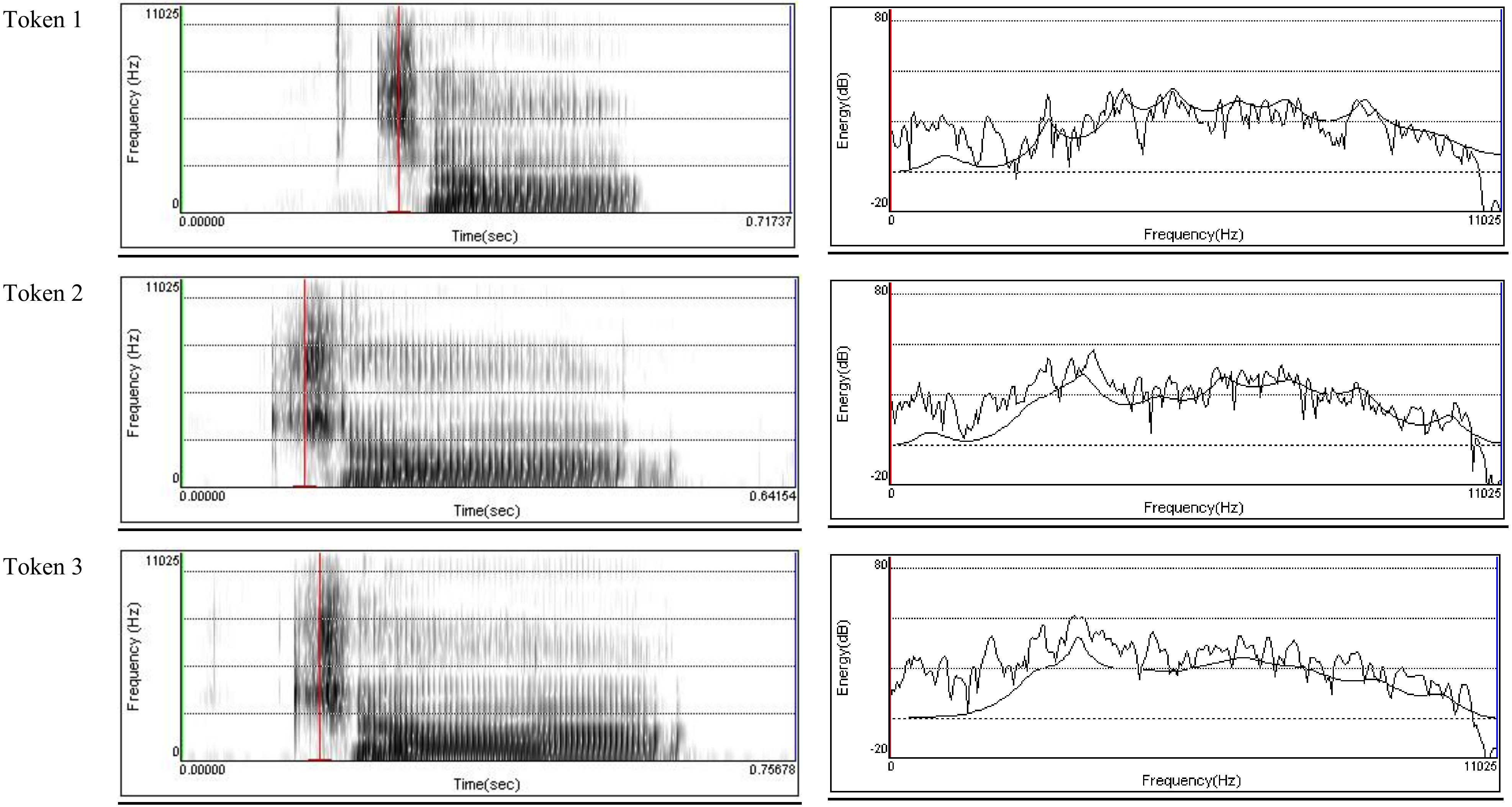


Figure 3.2.8. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [tʂ] in [tʂa˥] 渣 ‘residue’

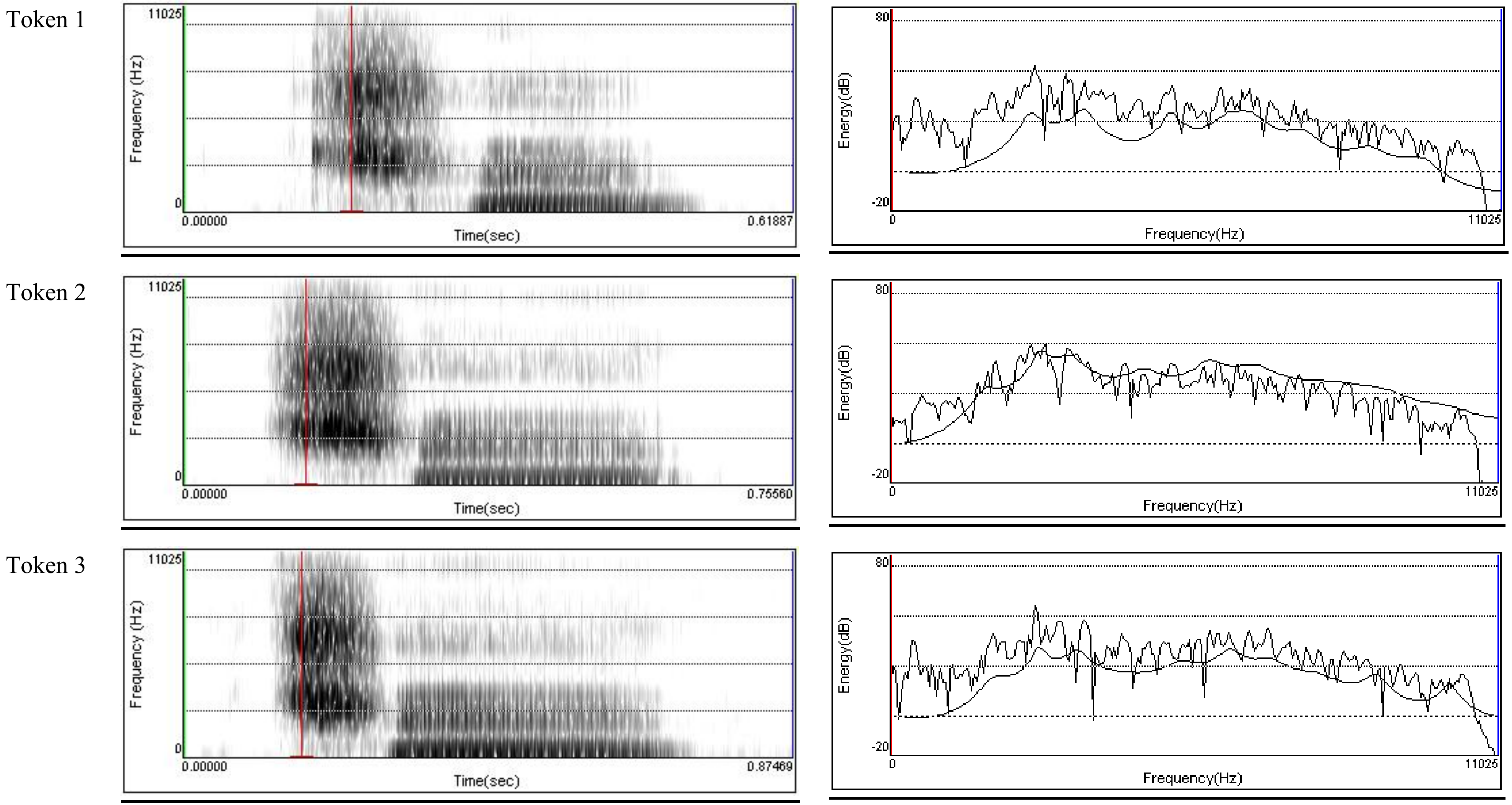


Figure 3.2.9. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [tʂʰ] in [tʂʰʅ˥] 吃 'eat'

Token 1

Token 2

Token 3

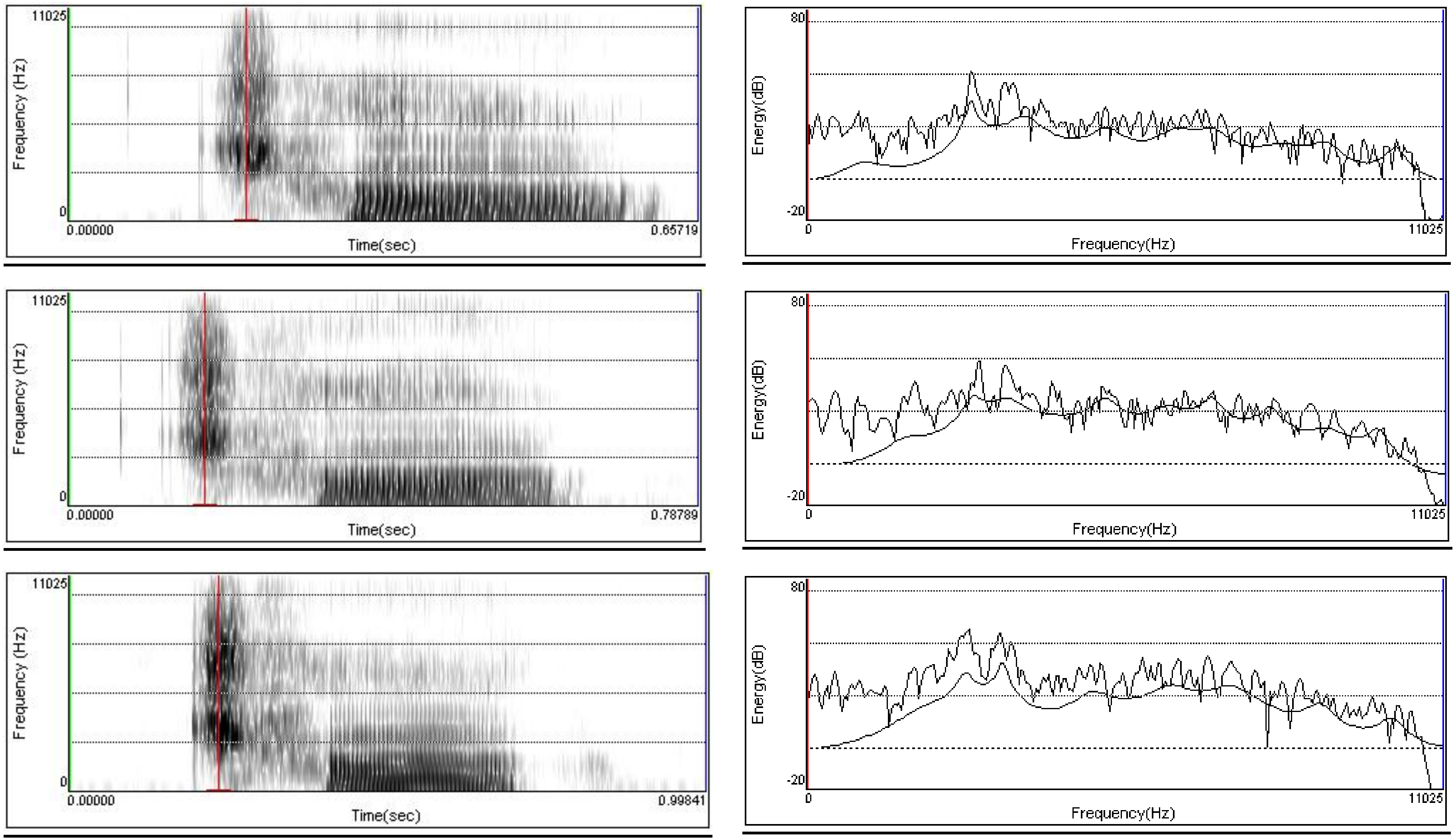


Figure 3.2.10. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [tʂ^h] in [tʂ^{h}a˥] 叉 ‘folk’

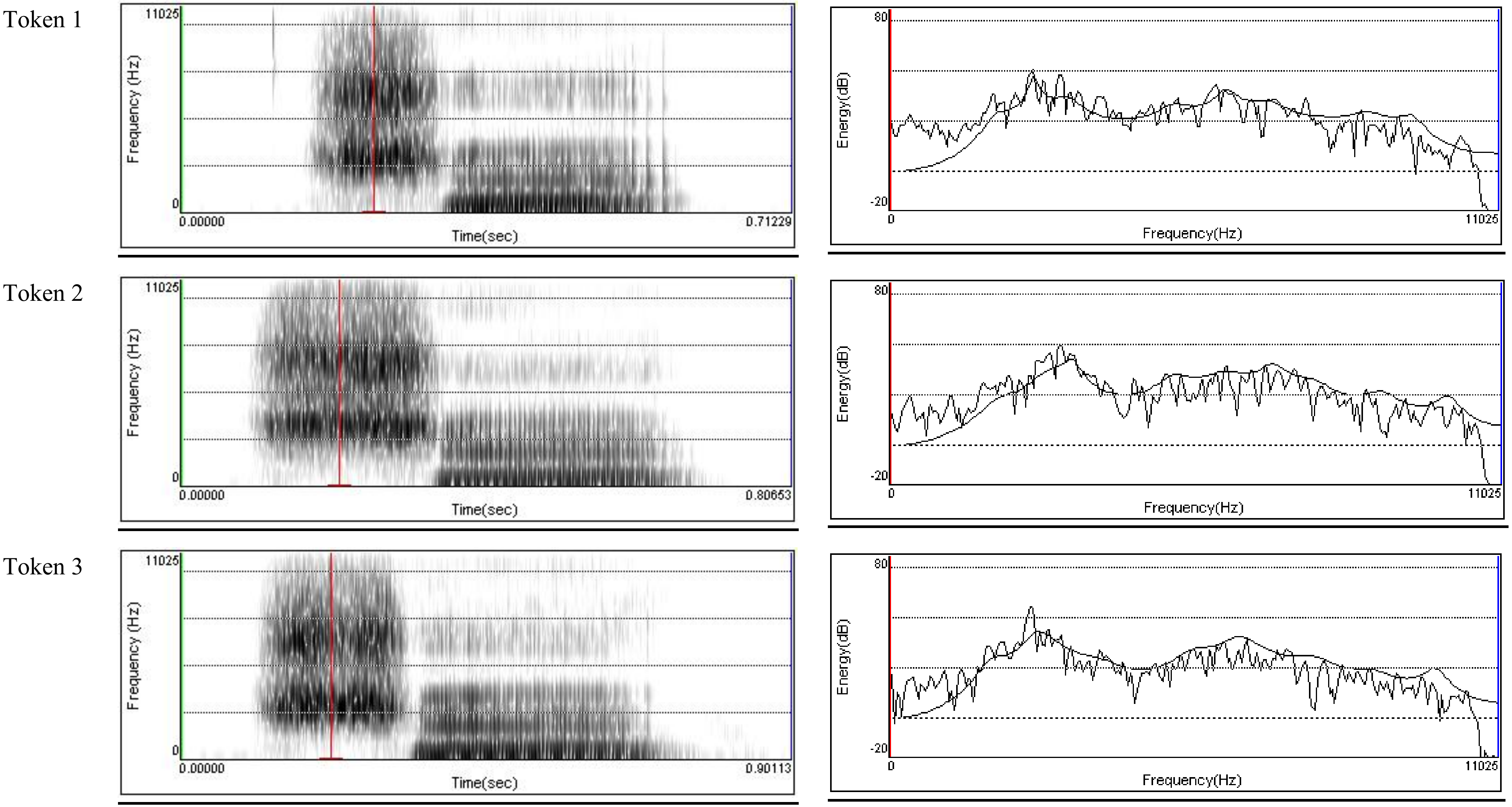


Figure 3.2.11. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [ʂ] in [ʂʅ˥] 失 'lose'

Token 1

Token 2

Token 3

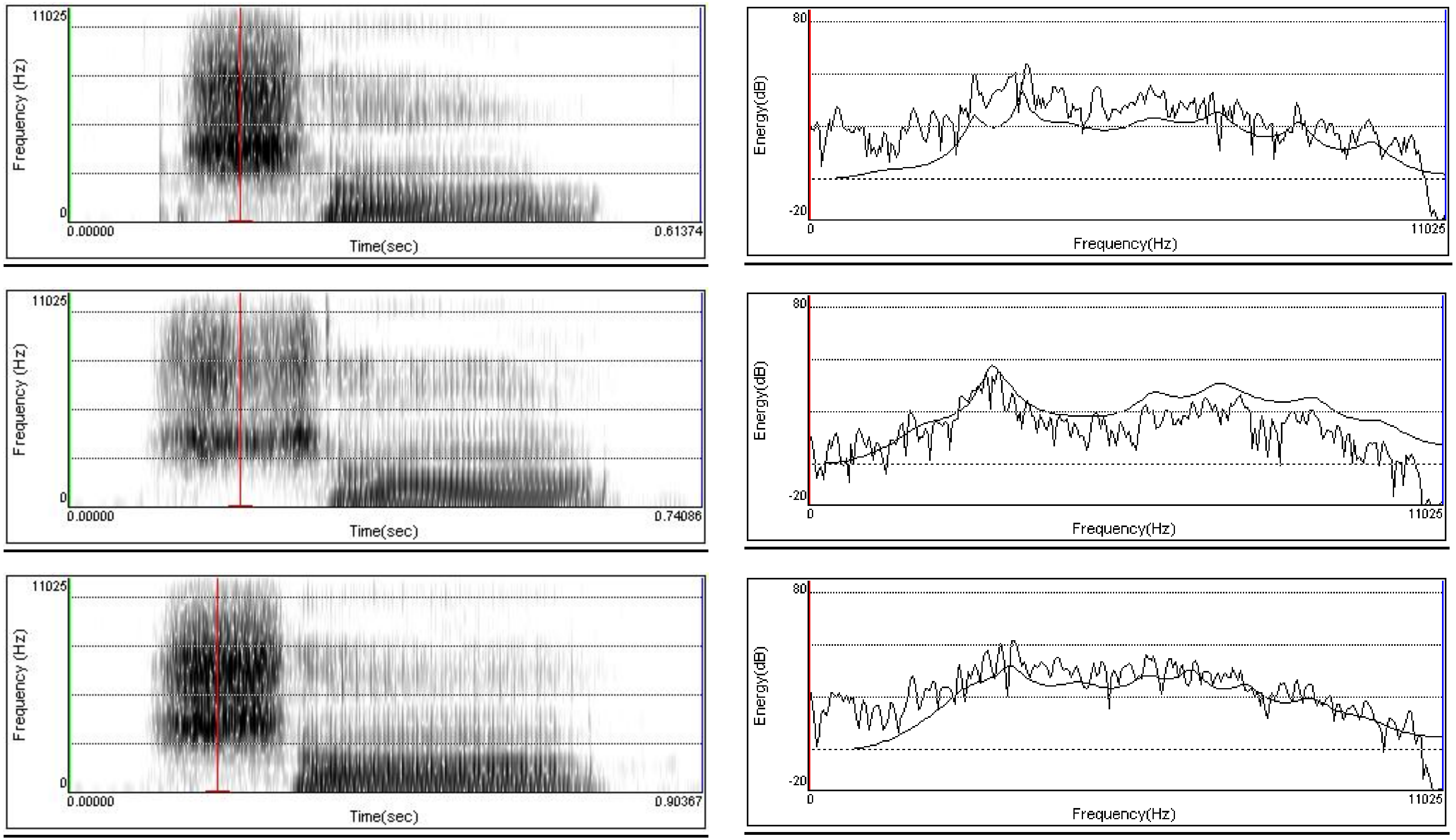


Figure 3.2.12. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [ʂ] in [ʂa˥] 沙 ‘sand’

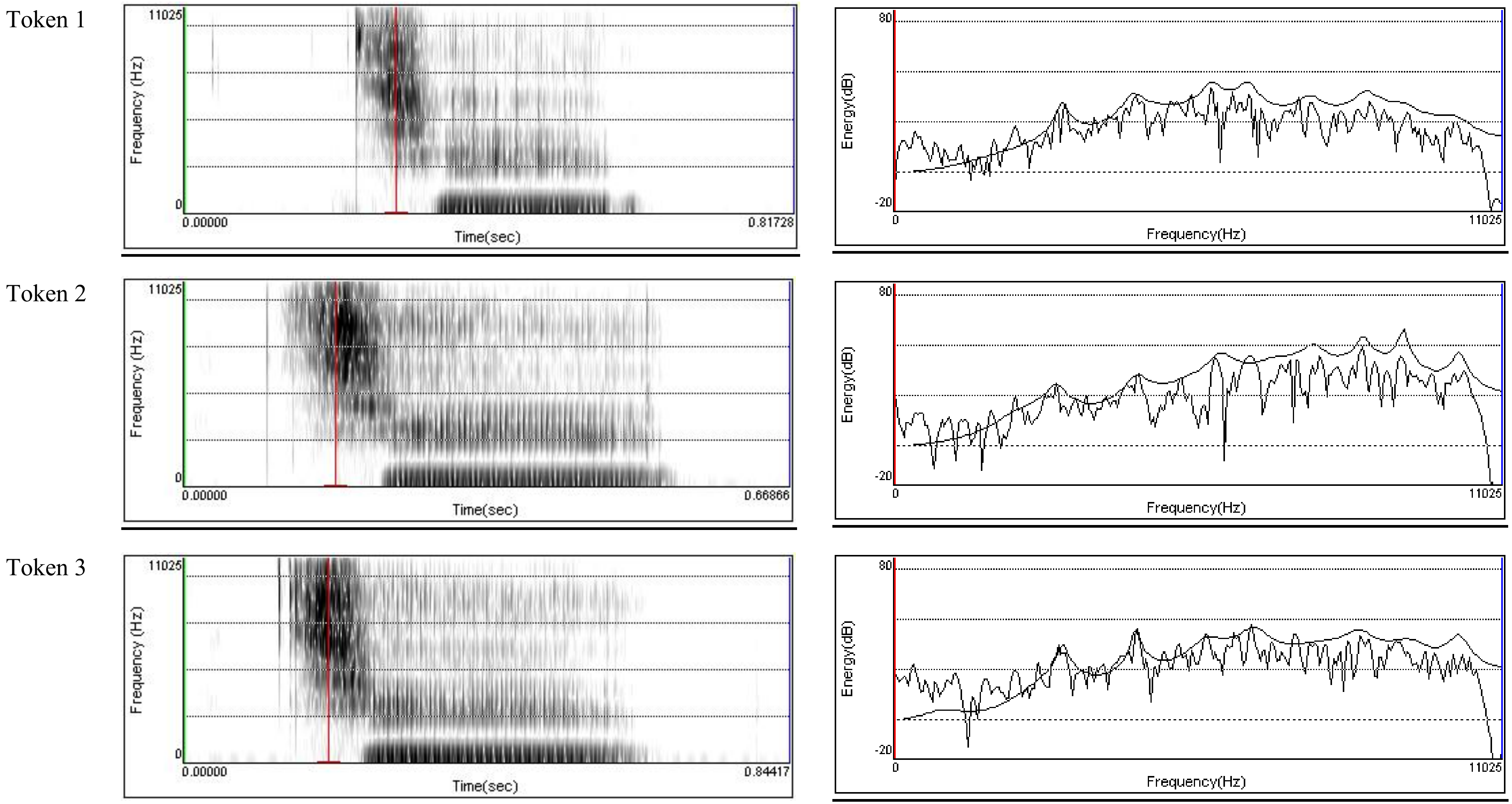


Figure 3.2.13. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [tɕ] in [tɕi˥] 基 ‘base’

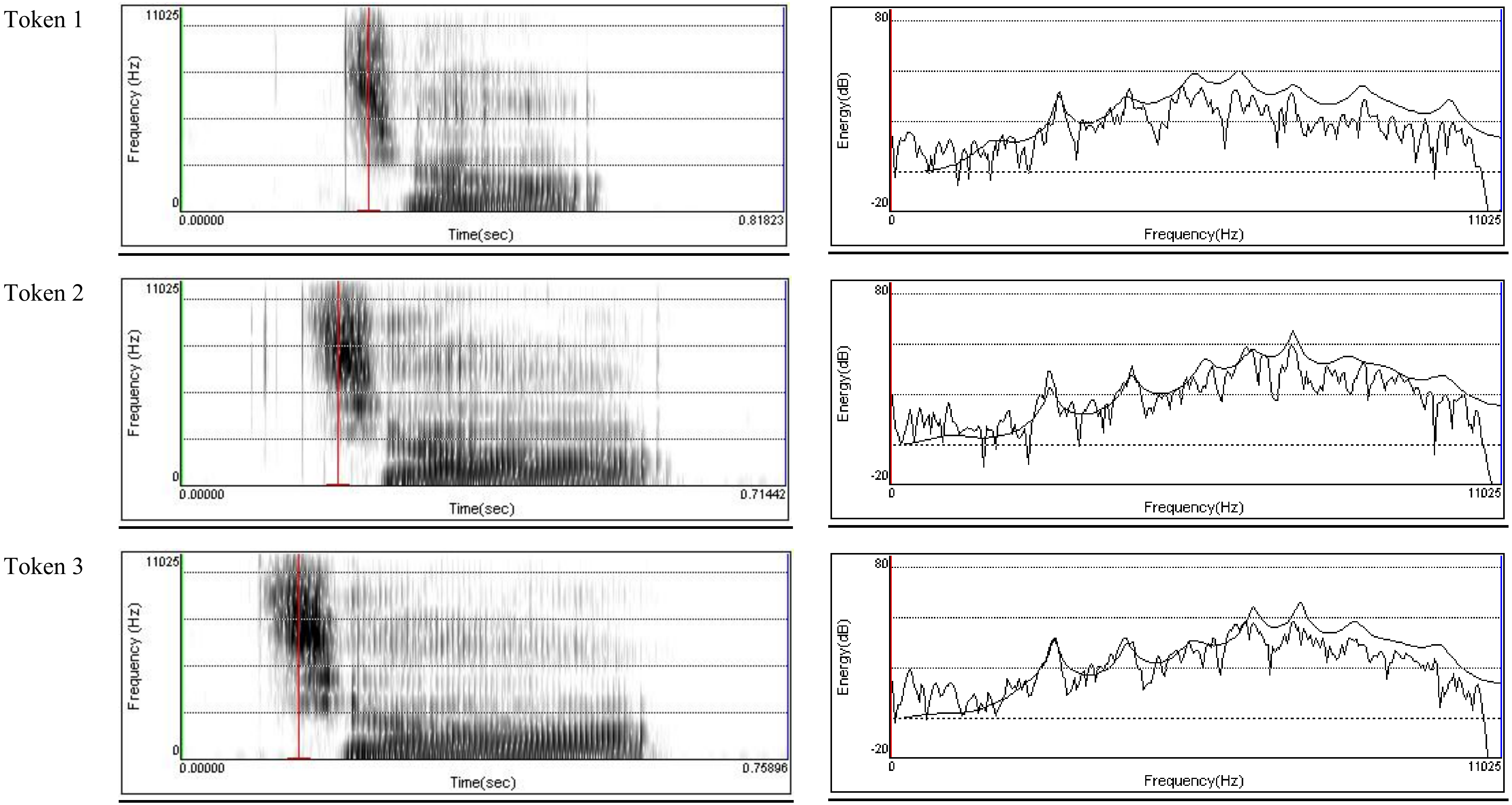


Figure 3.2.14. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [tɕ] in [tɕia˥] 加 ‘add’

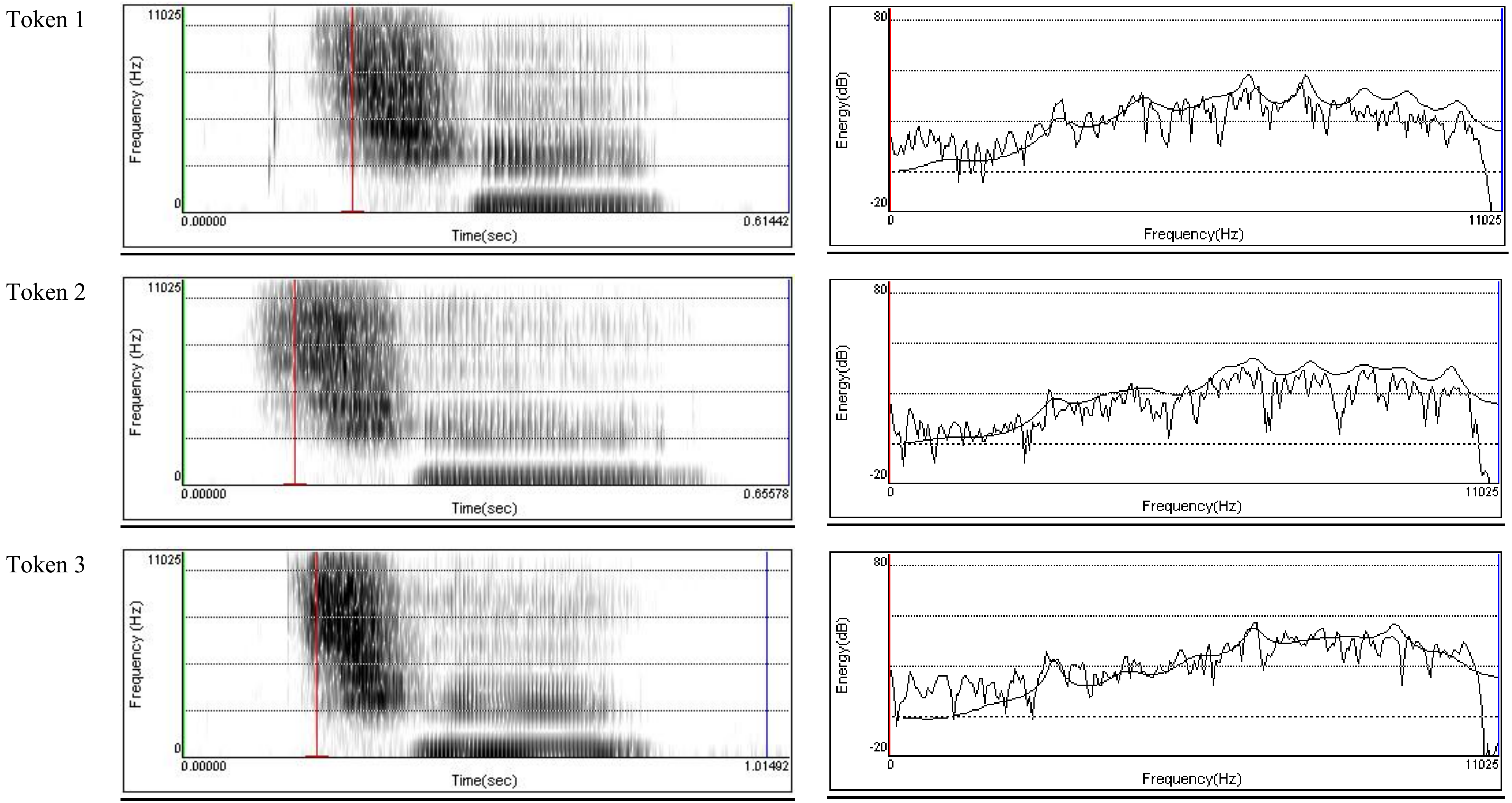


Figure 3.2.15. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [tɕʰ] in [tɕʰi˥] 七 'seven'

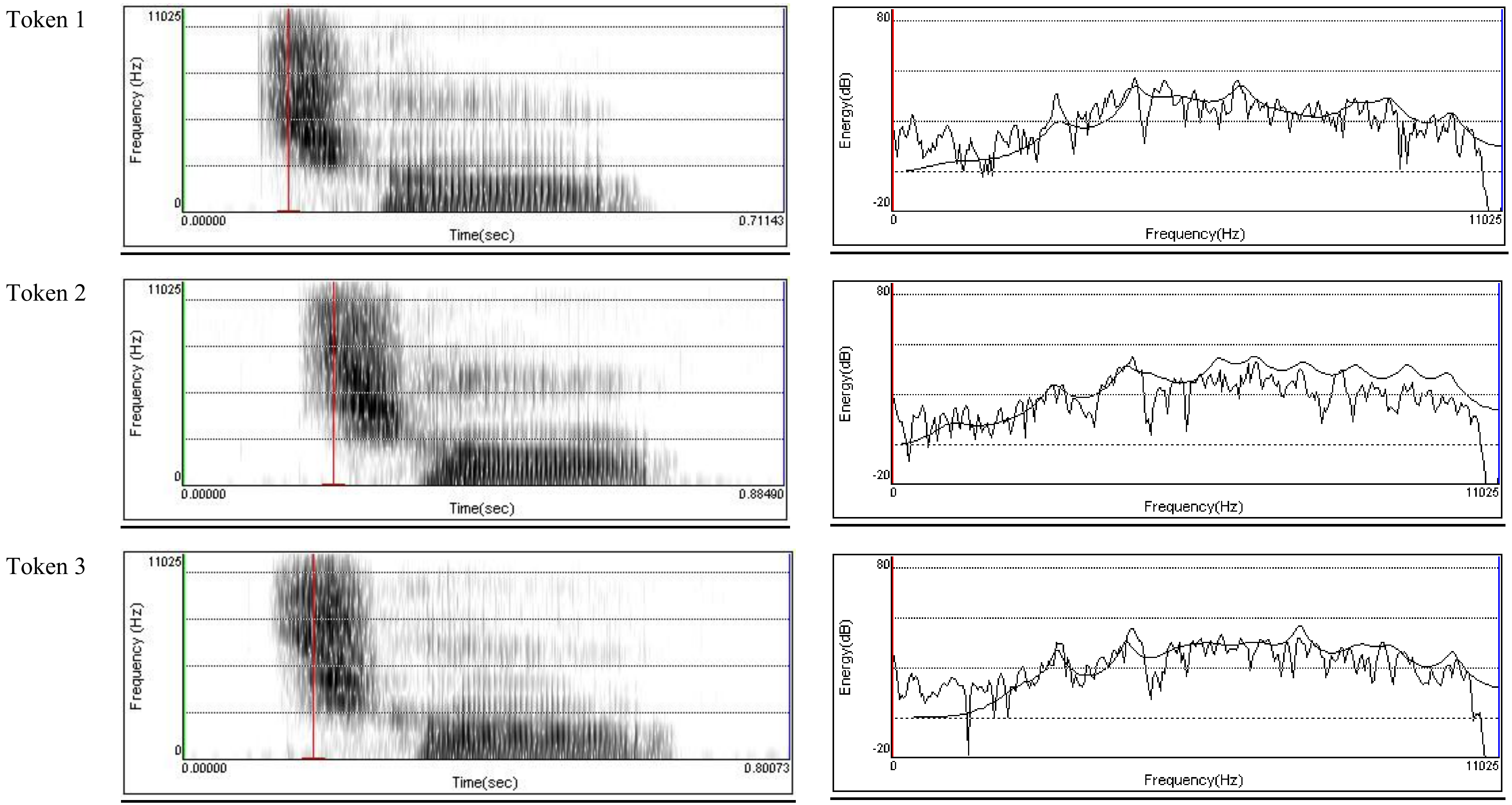


Figure 3.2.16. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [tɕʰ] in [tɕʰia˥] 掐 ‘nip off’

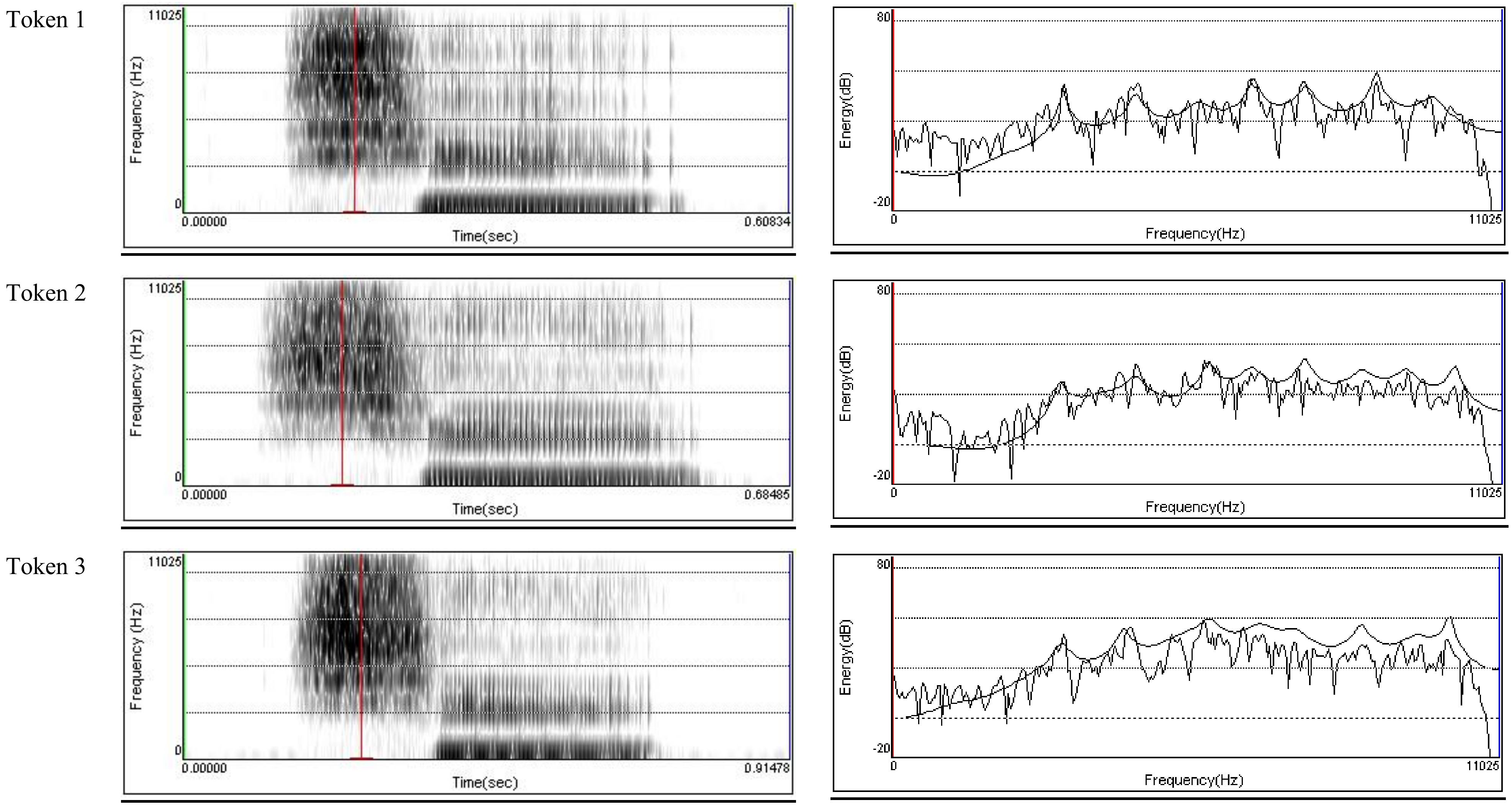


Figure 3.2.17. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [ɕ] in [ɕi˥] 希 'hope'

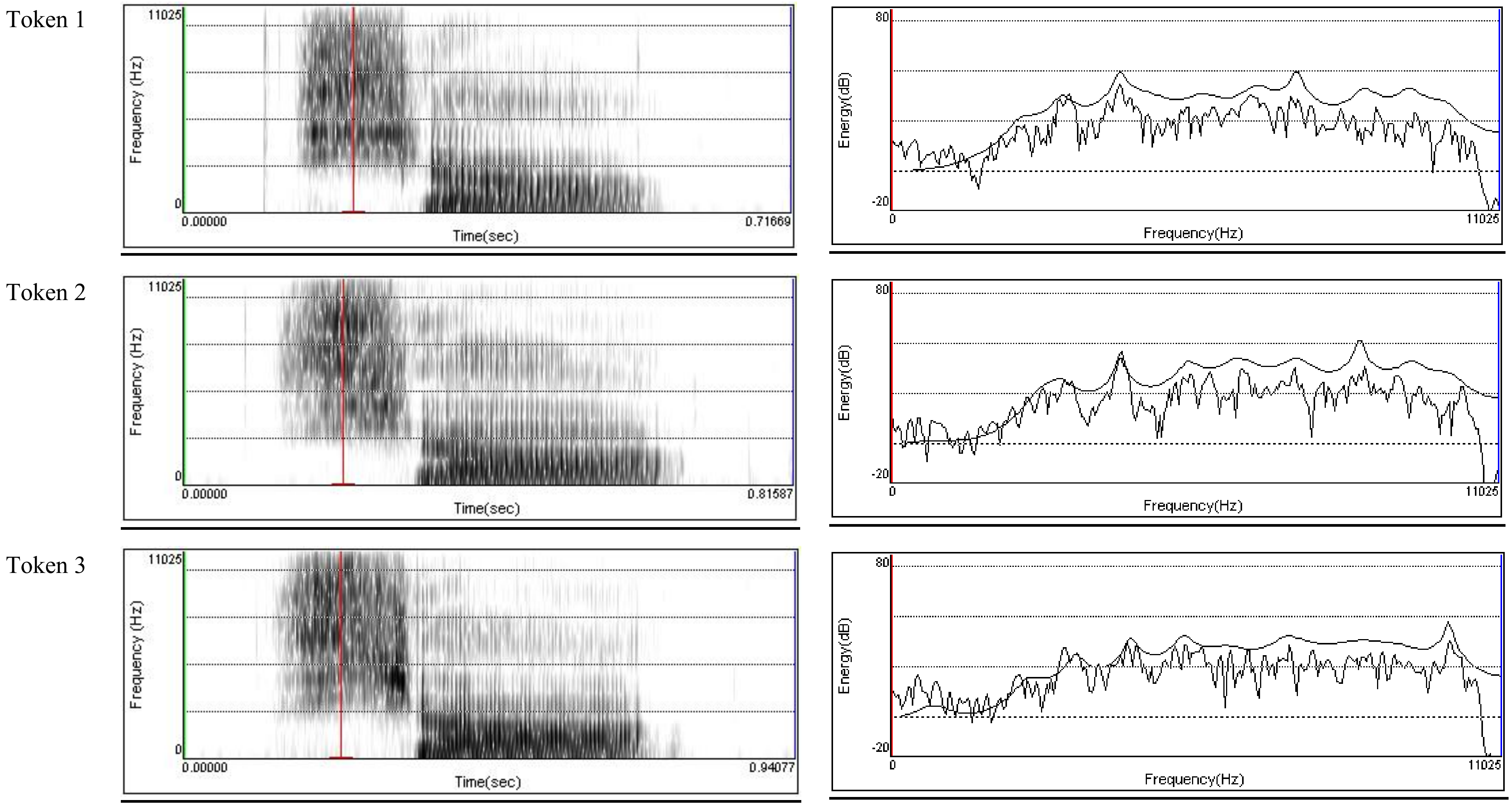


Figure 3.2.18. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [ɕ] in [ɕia˥] 蝦 ‘prawn’

3.3 Cantonese Male 2

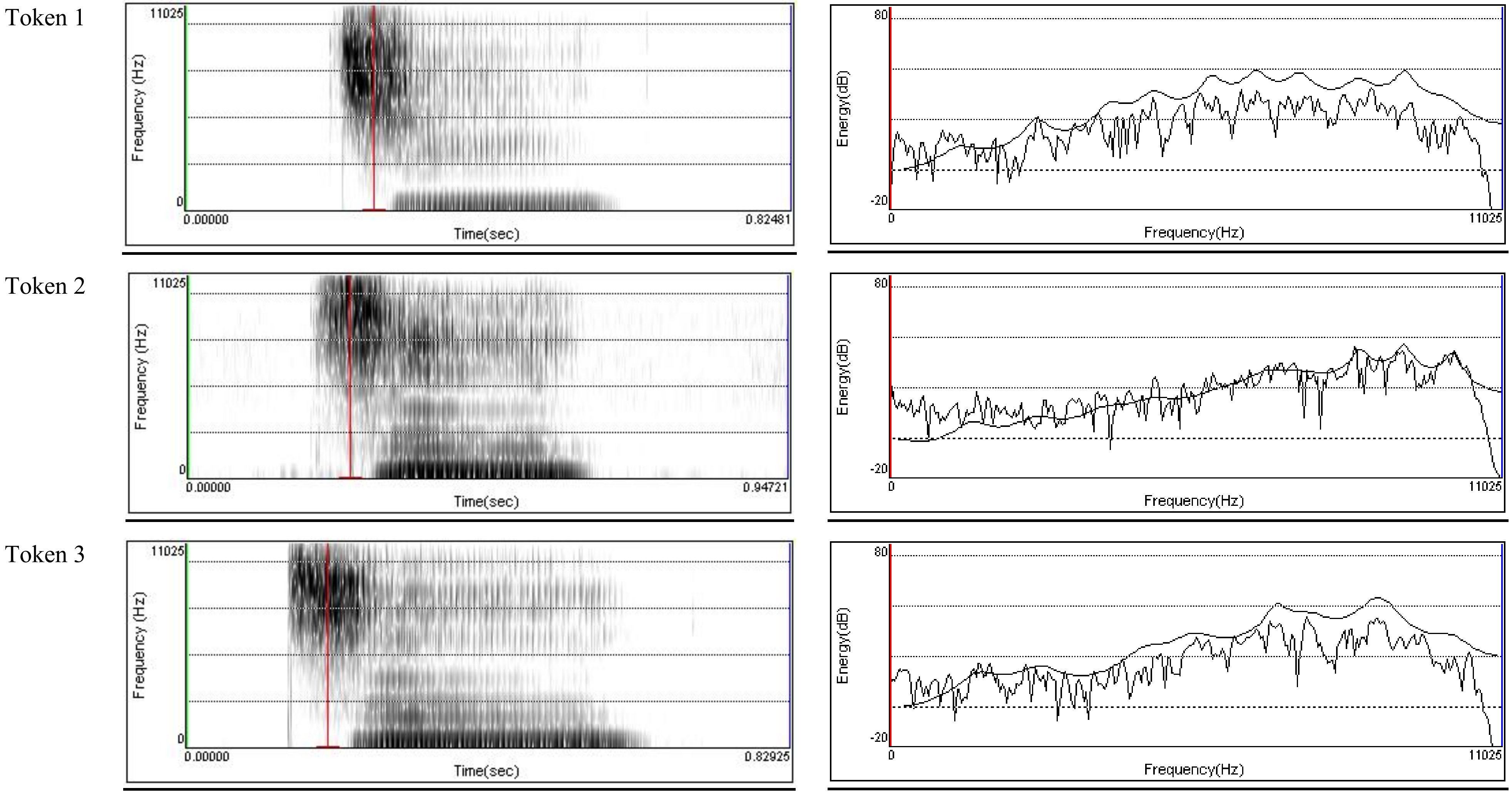


Figure 3.3.1. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [ʦ] in [ʦɿ˥] 資 ‘capital’

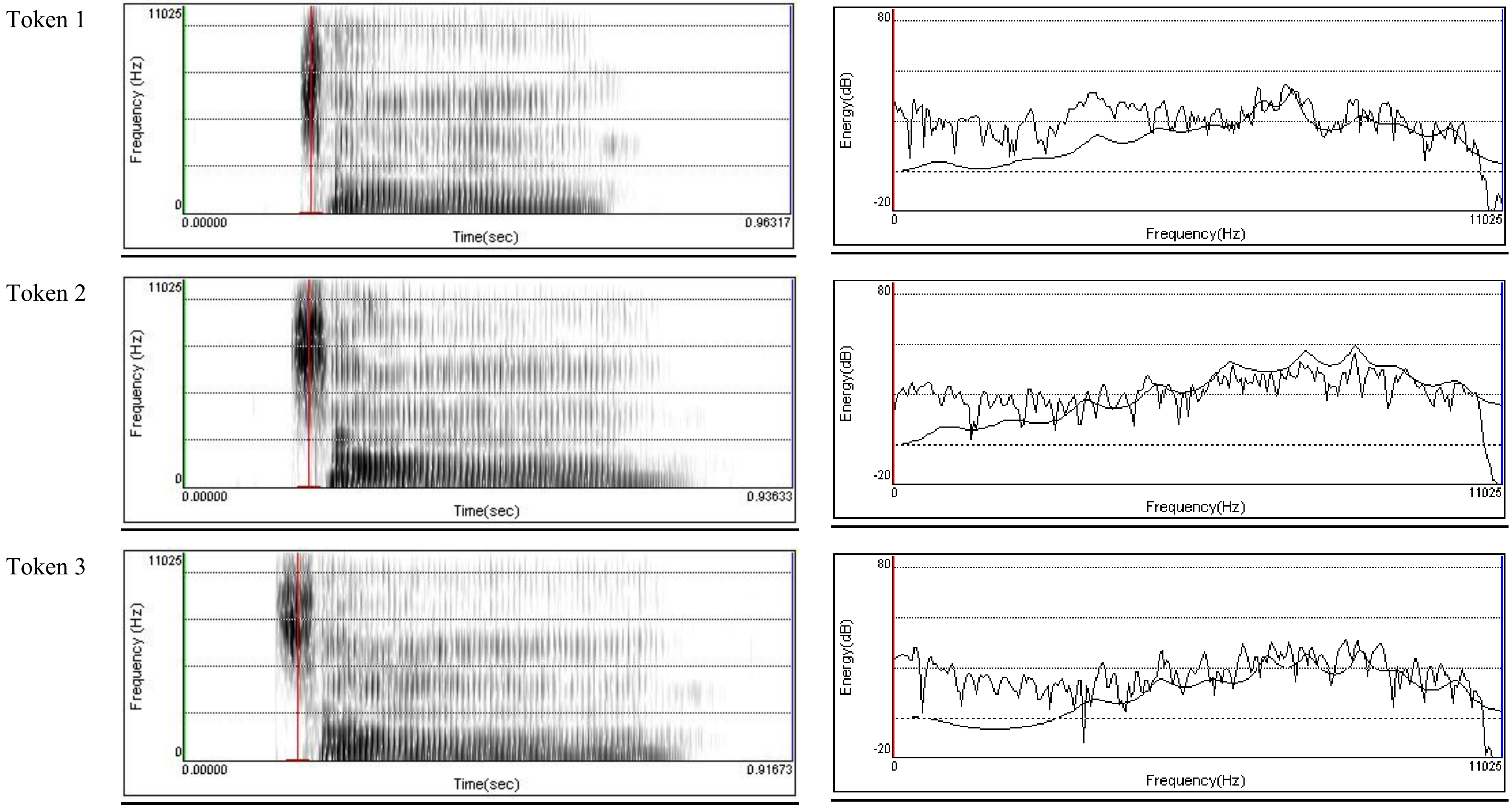


Figure 3.3.2. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [ʦ] in [ʦa˥] 紮 ‘tie’

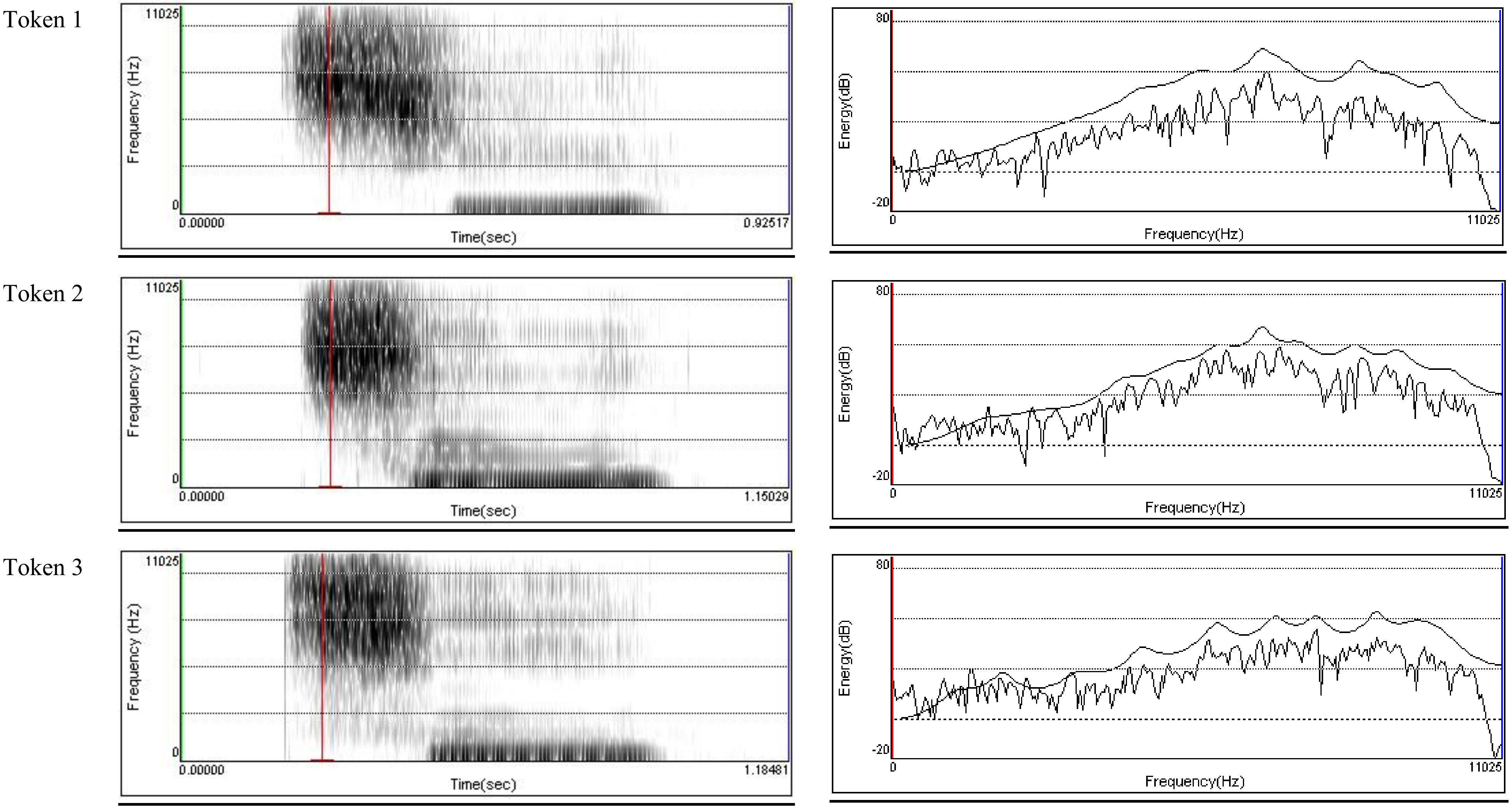


Figure 3.3.3. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [ʦʰ] in [ʦʰɿ˥] 疵 'flaw'

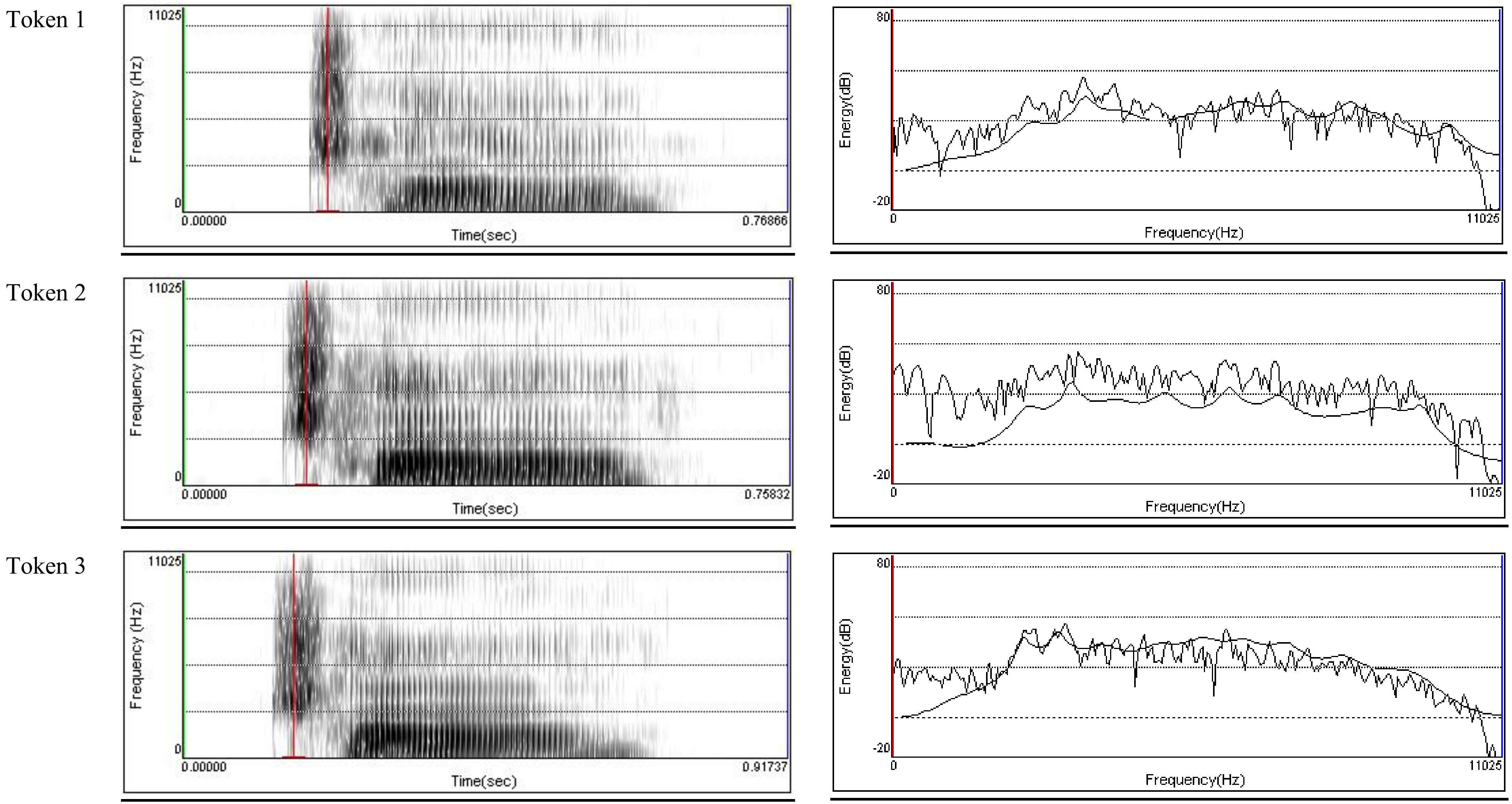


Figure 3.3.4. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [ʦʰ] in [ʦʰa˥] 擦 ‘rub’

Token 1

Token 2

Token 3

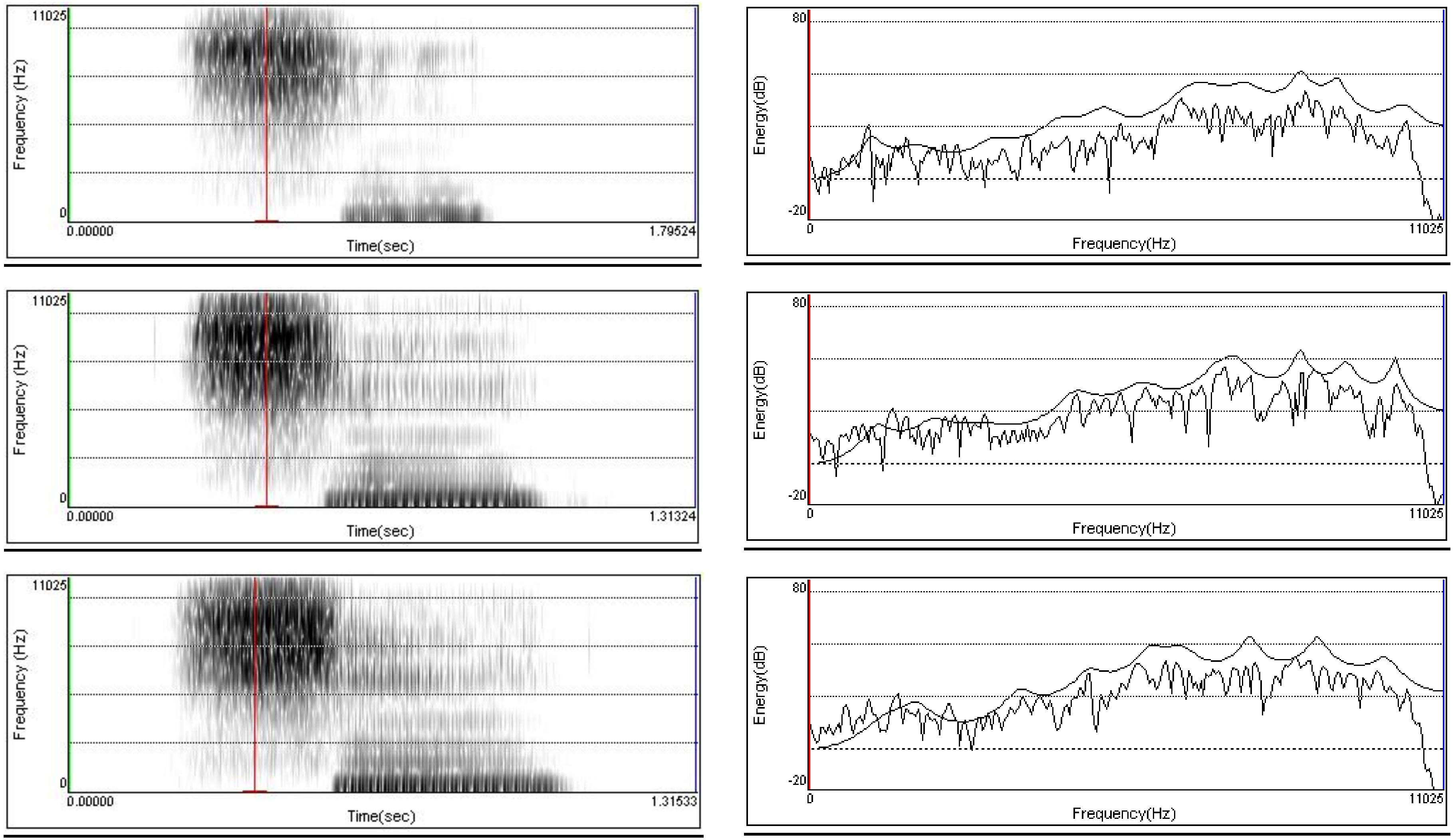


Figure 3.3.5. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [s] in [sɿ˥] 司 ‘in charge of’

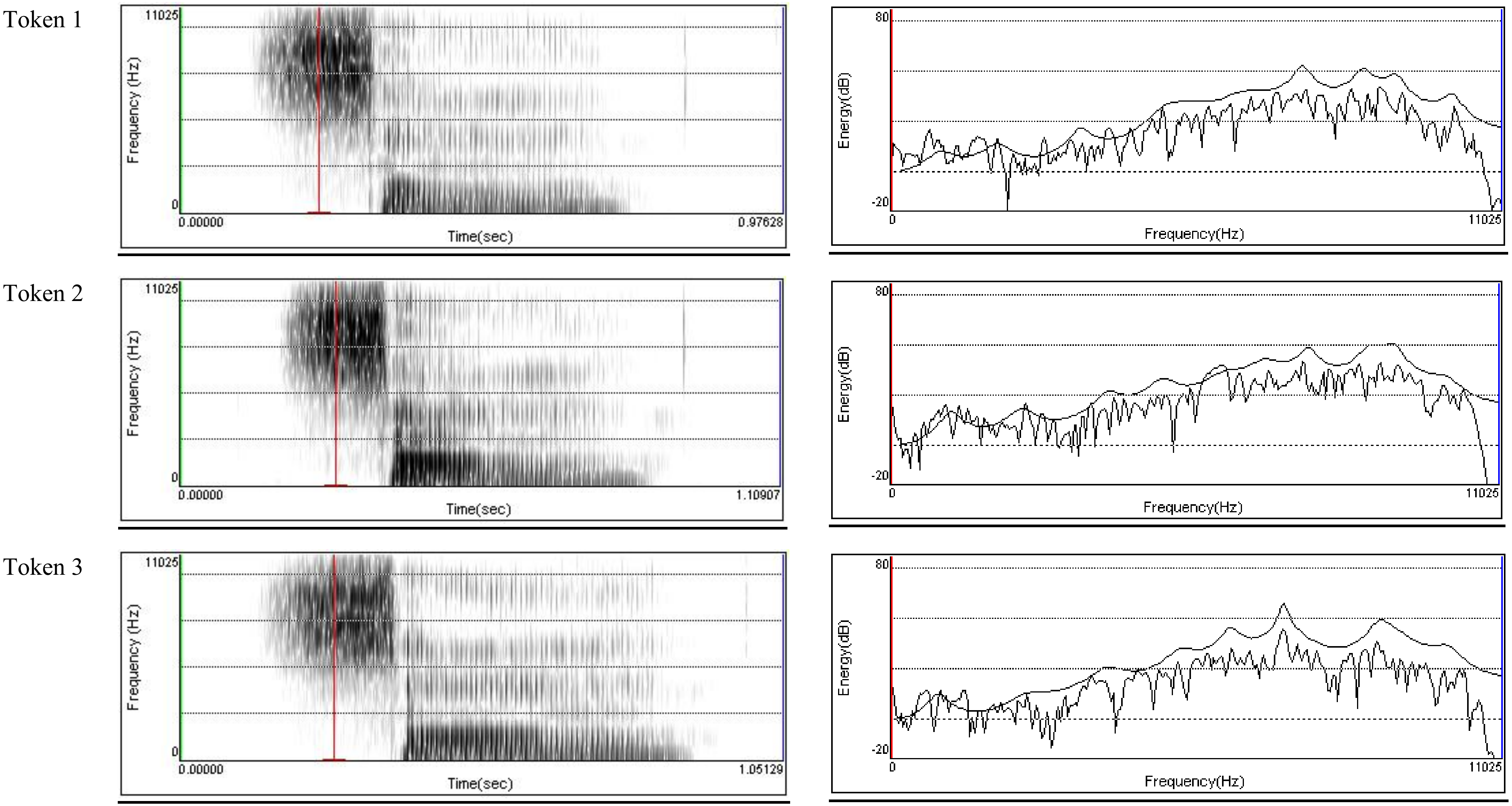


Figure 3.3.6. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [s] in [sa˥] 撒 'cast'

Token 1

Token 2

Token 3

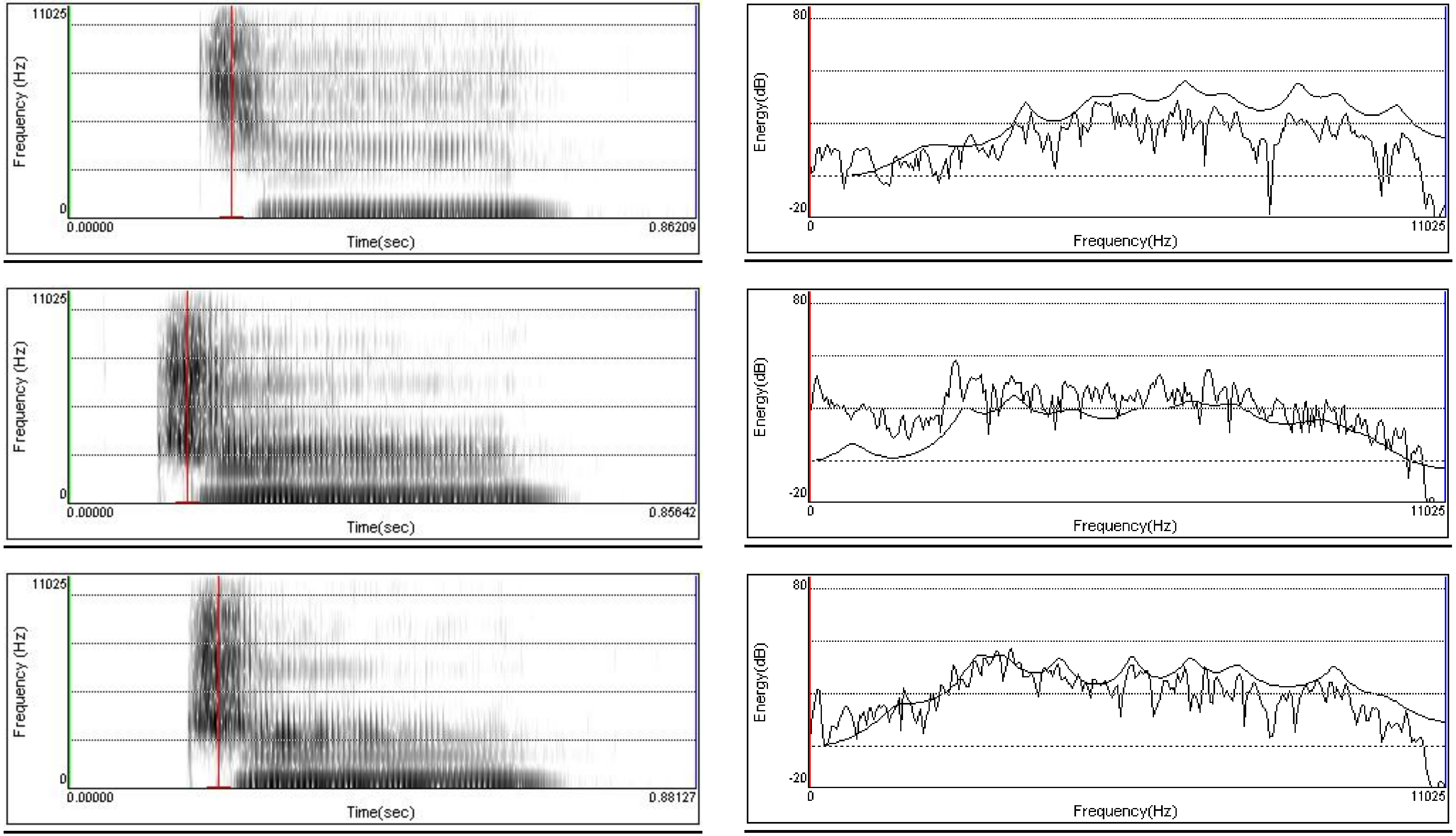


Figure 3.3.7. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [tʂ] in [tʂʅ˥] 知 ‘know’

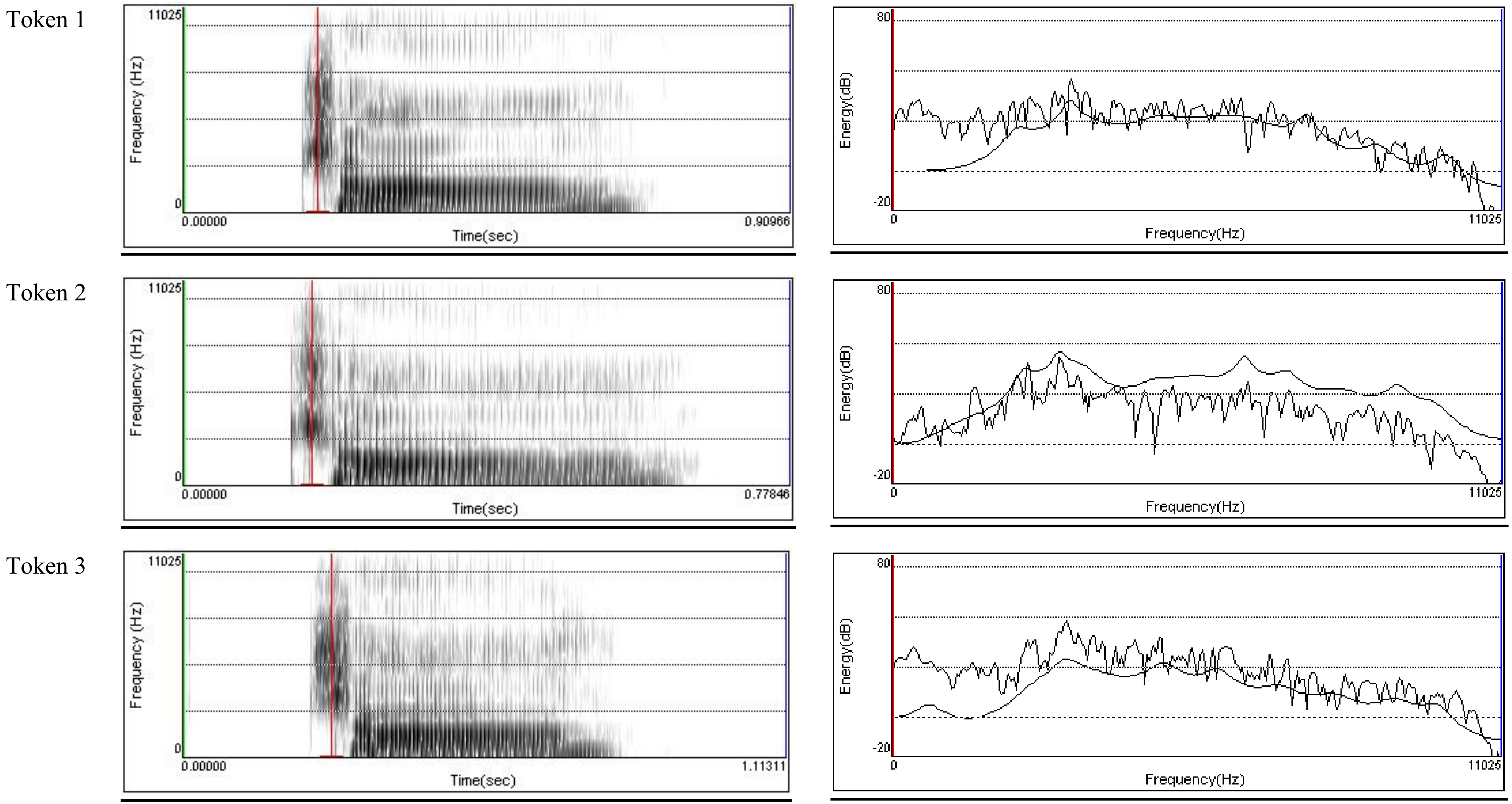


Figure 3.3.8. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [tʂ] in [tʂa˥] 渣 ‘residue’

Token 1

Token 2

Token 3

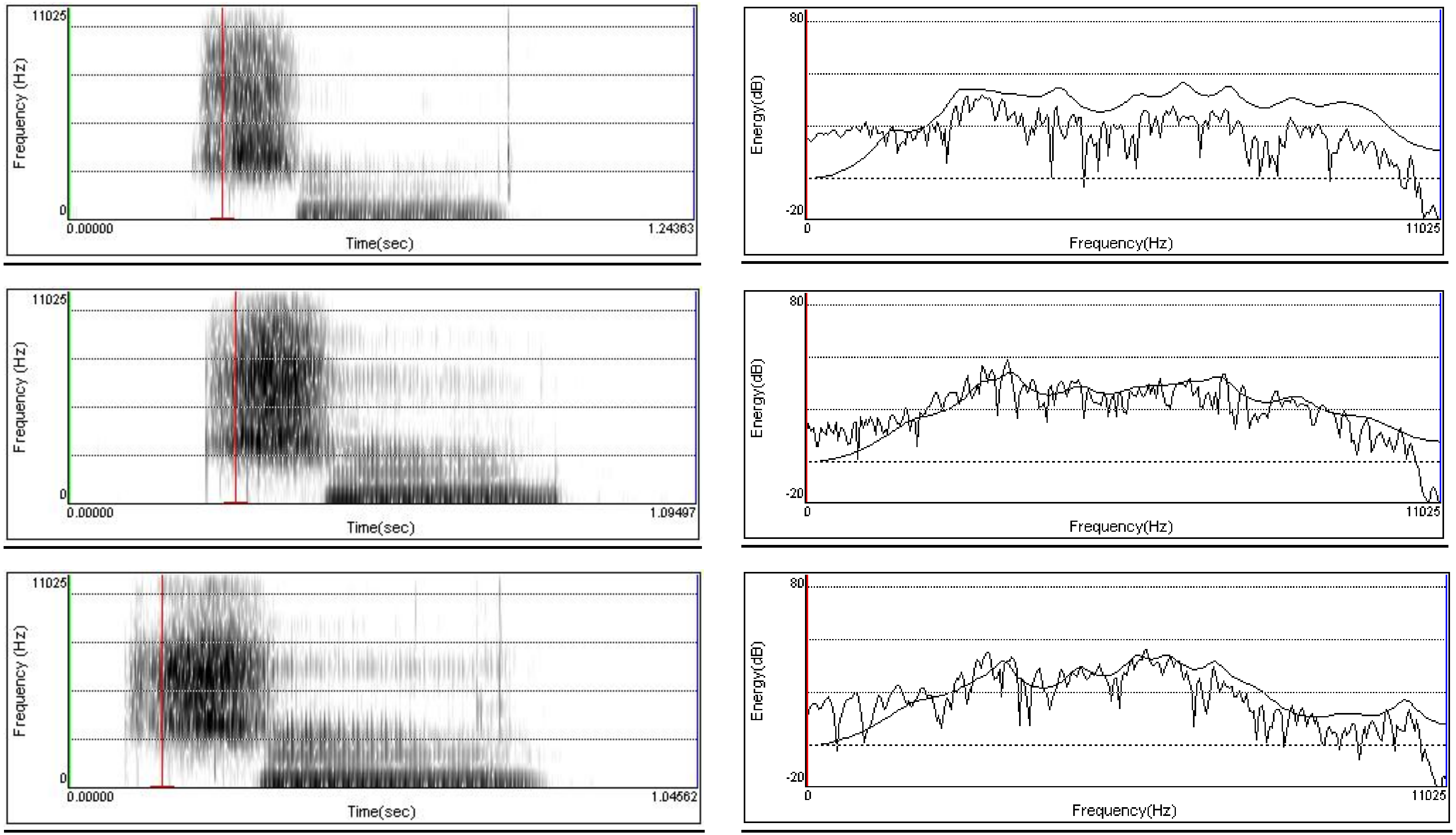


Figure 3.3.9. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [tʂʰ] in [tʂʰʅ˥] 吃 'eat'

Token 1

Token 2

Token 3

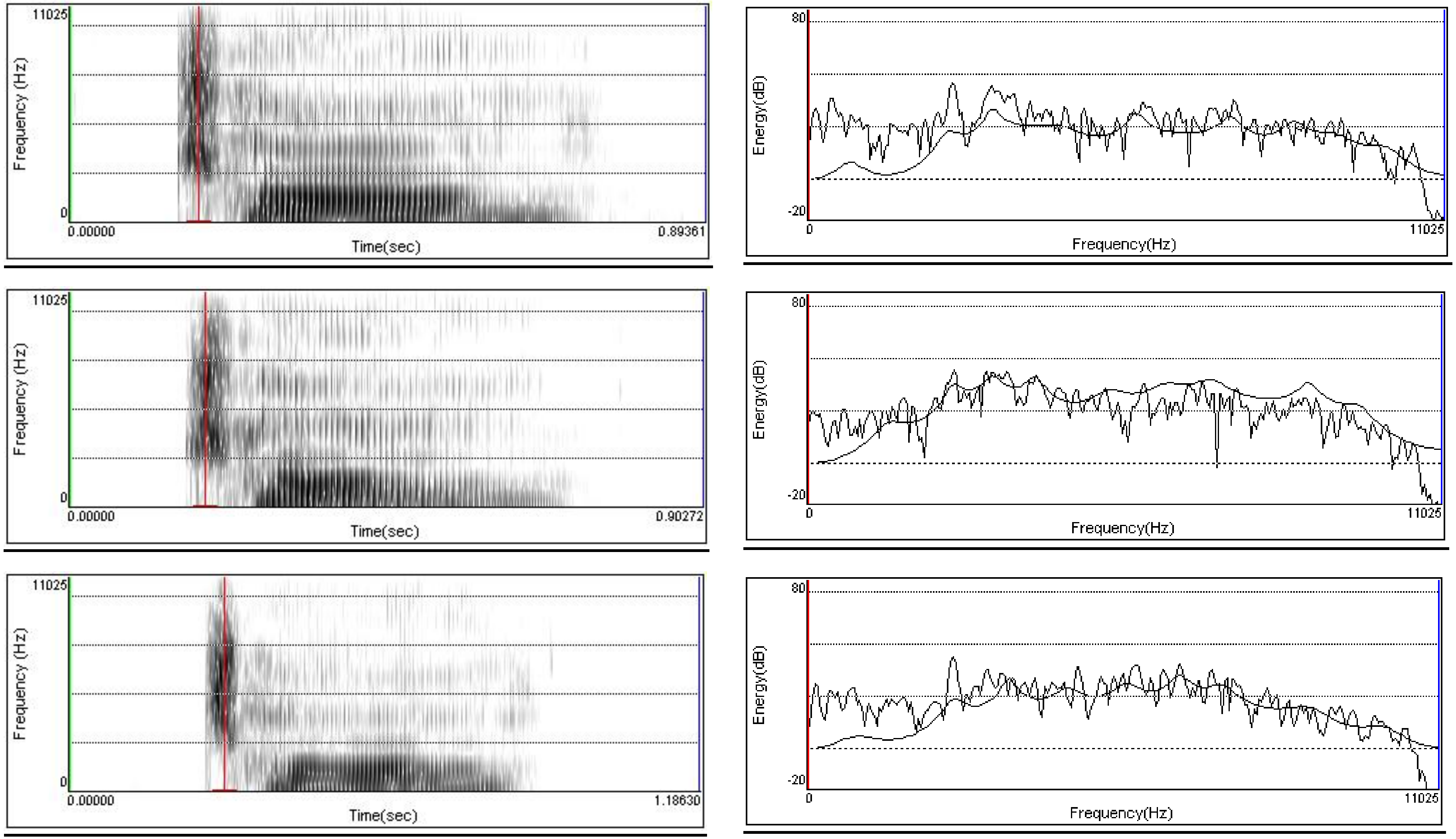


Figure 3.3.10. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [tʂ^{h}] in [tʂ^{h}a˥] 叉 'folk'

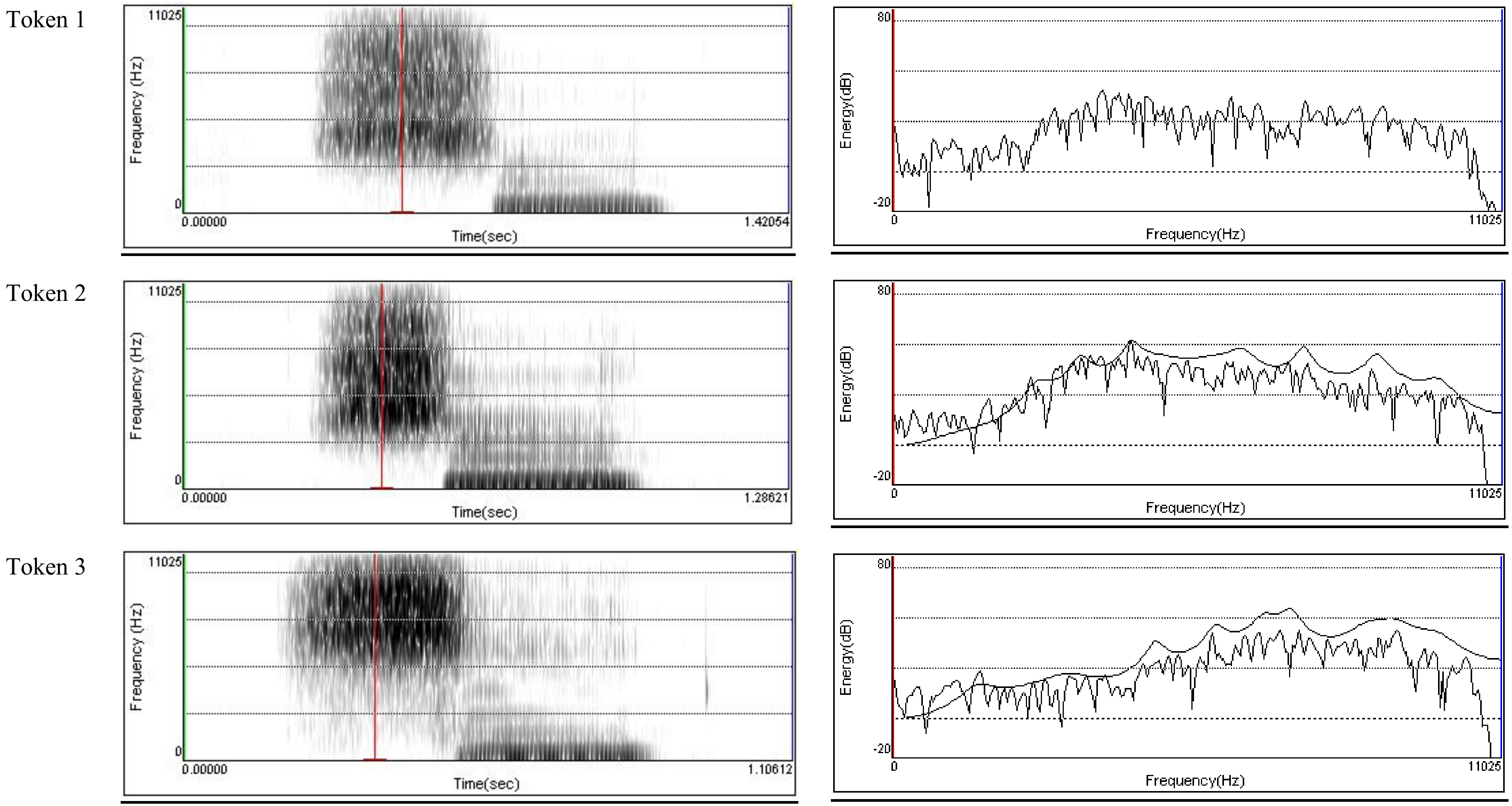


Figure 3.3.11. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [ʂ] in [ʂʅ˥] 失 ‘lose’

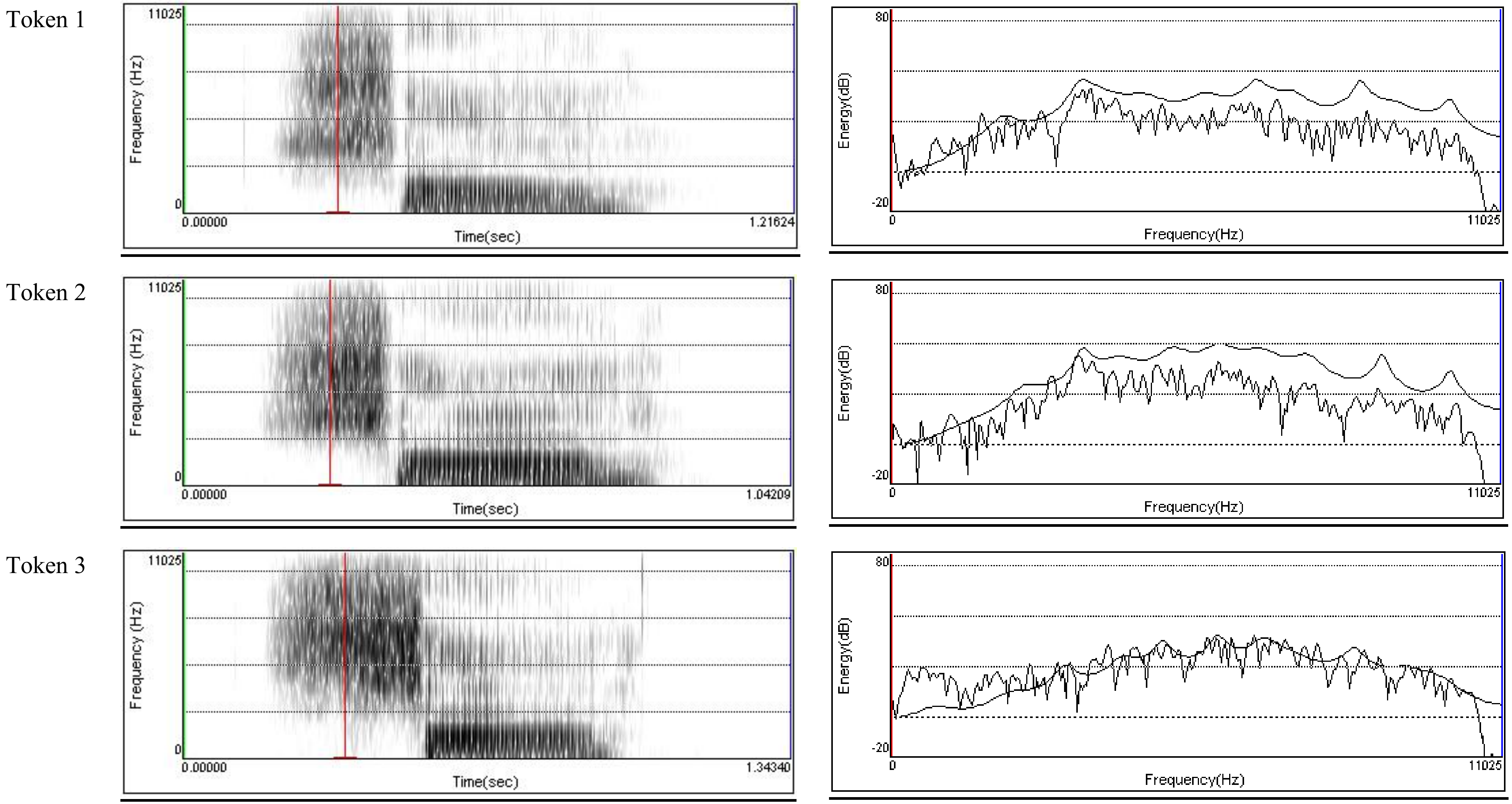


Figure 3.3.12. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [ʂ] in [ʂa˥] 沙 'sand'

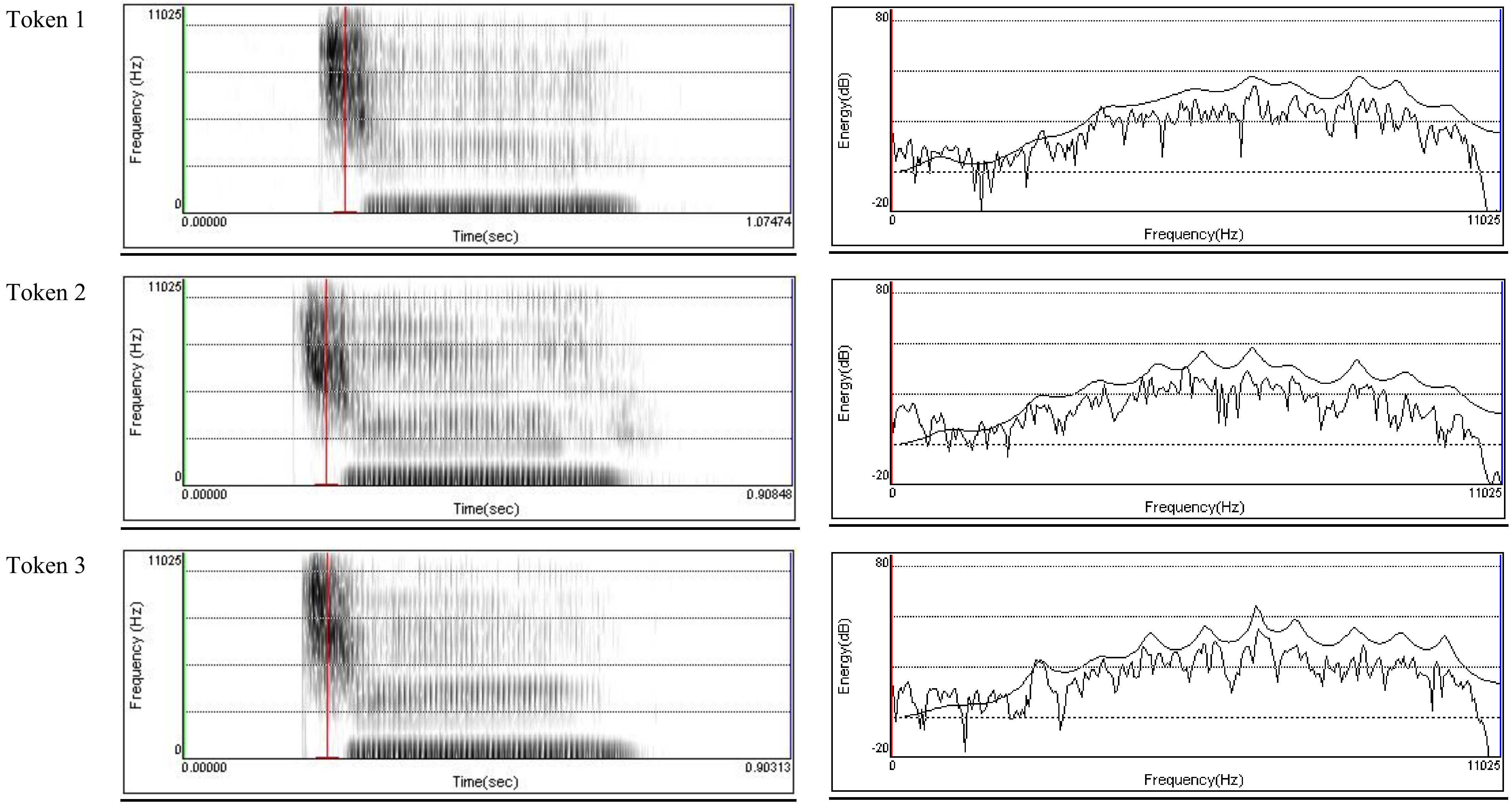


Figure 3.3.13. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [tɕ] in [tɕi˥] 基 ‘base’

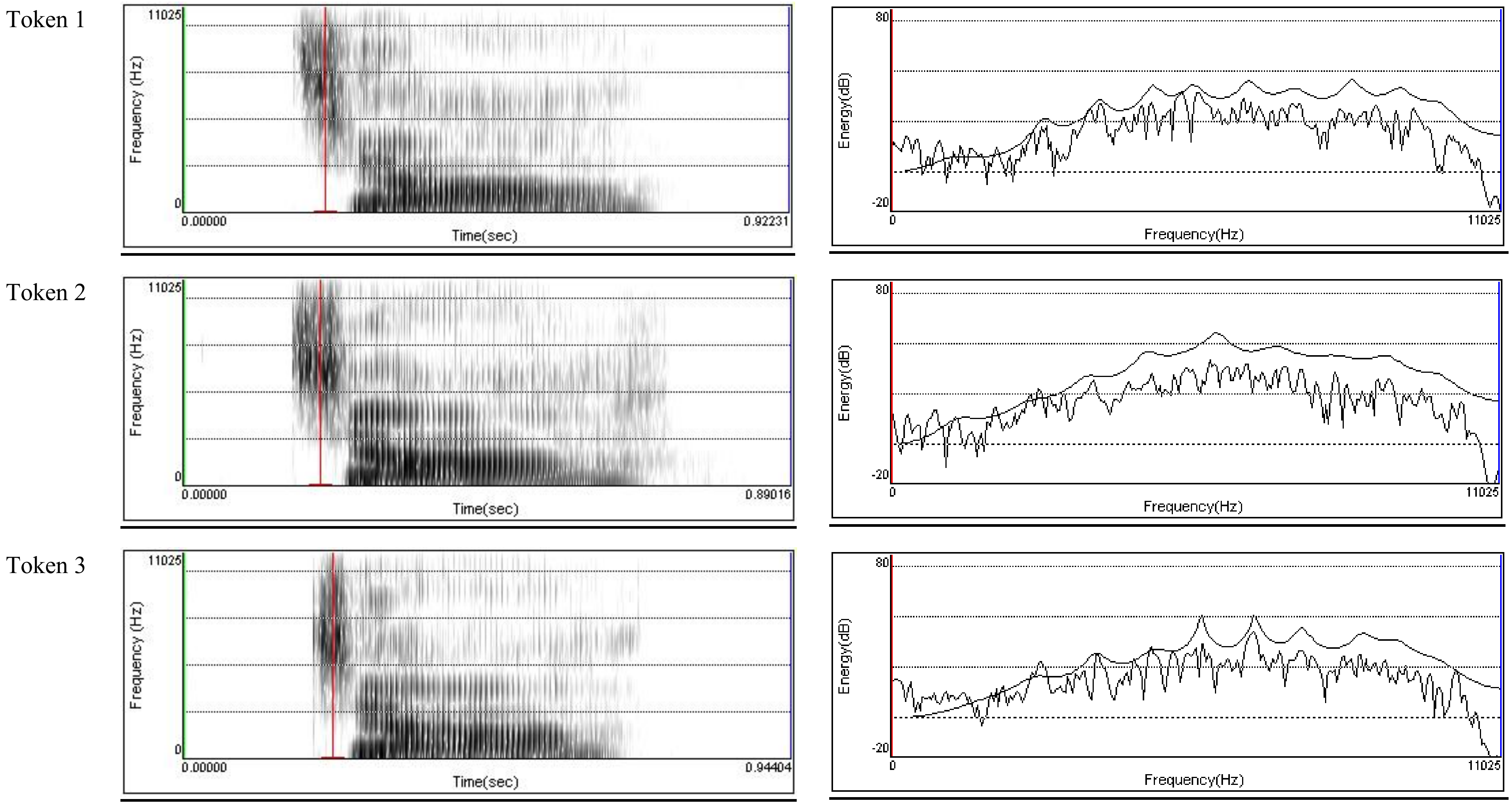


Figure 3.3.14. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [tɕ] in [tɕia˥] 加 ‘add’

Token 1

Token 2

Token 3

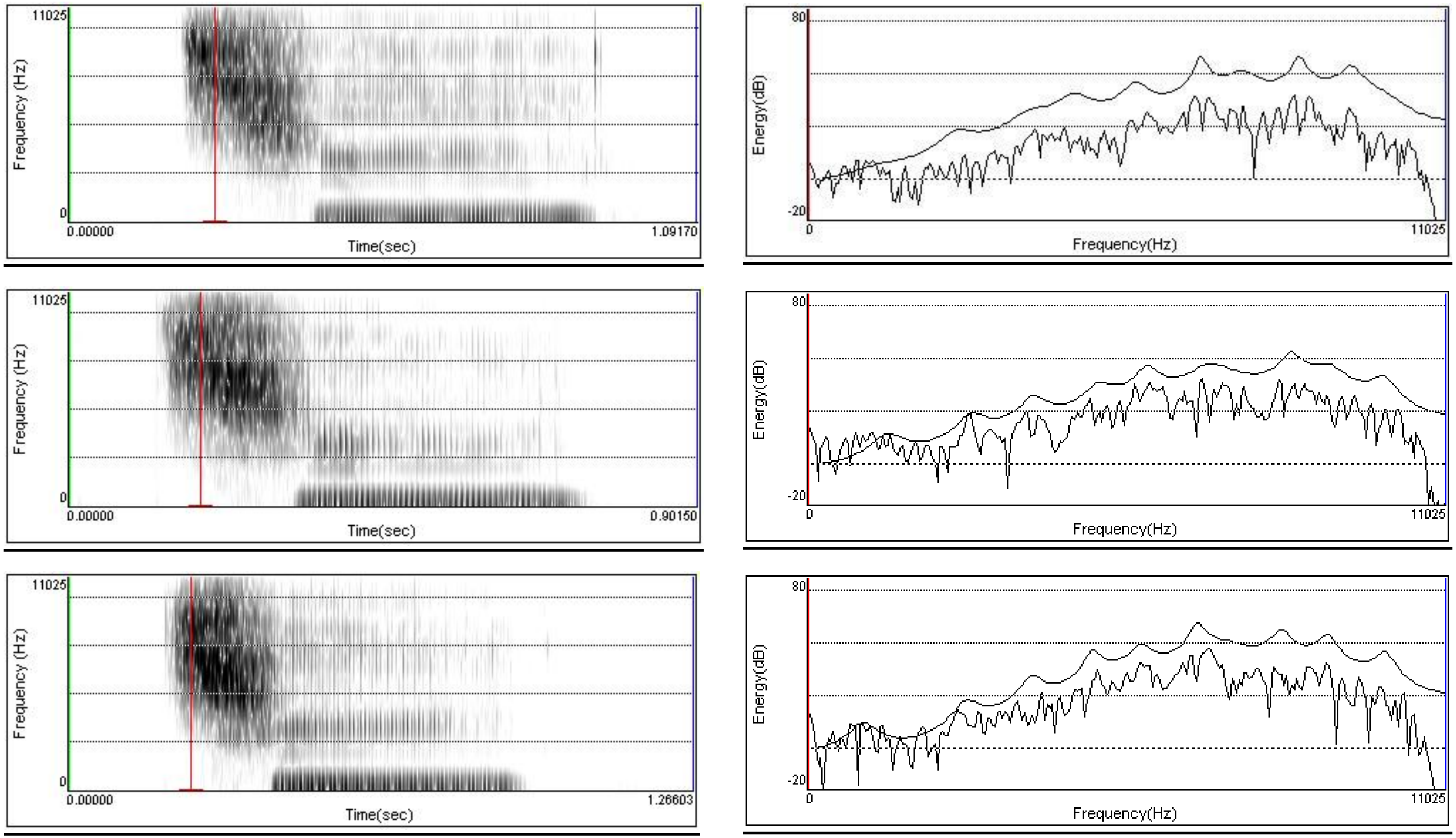


Figure 3.3.15. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [tɕʰ] in [tɕʰi˥] 七 'seven'

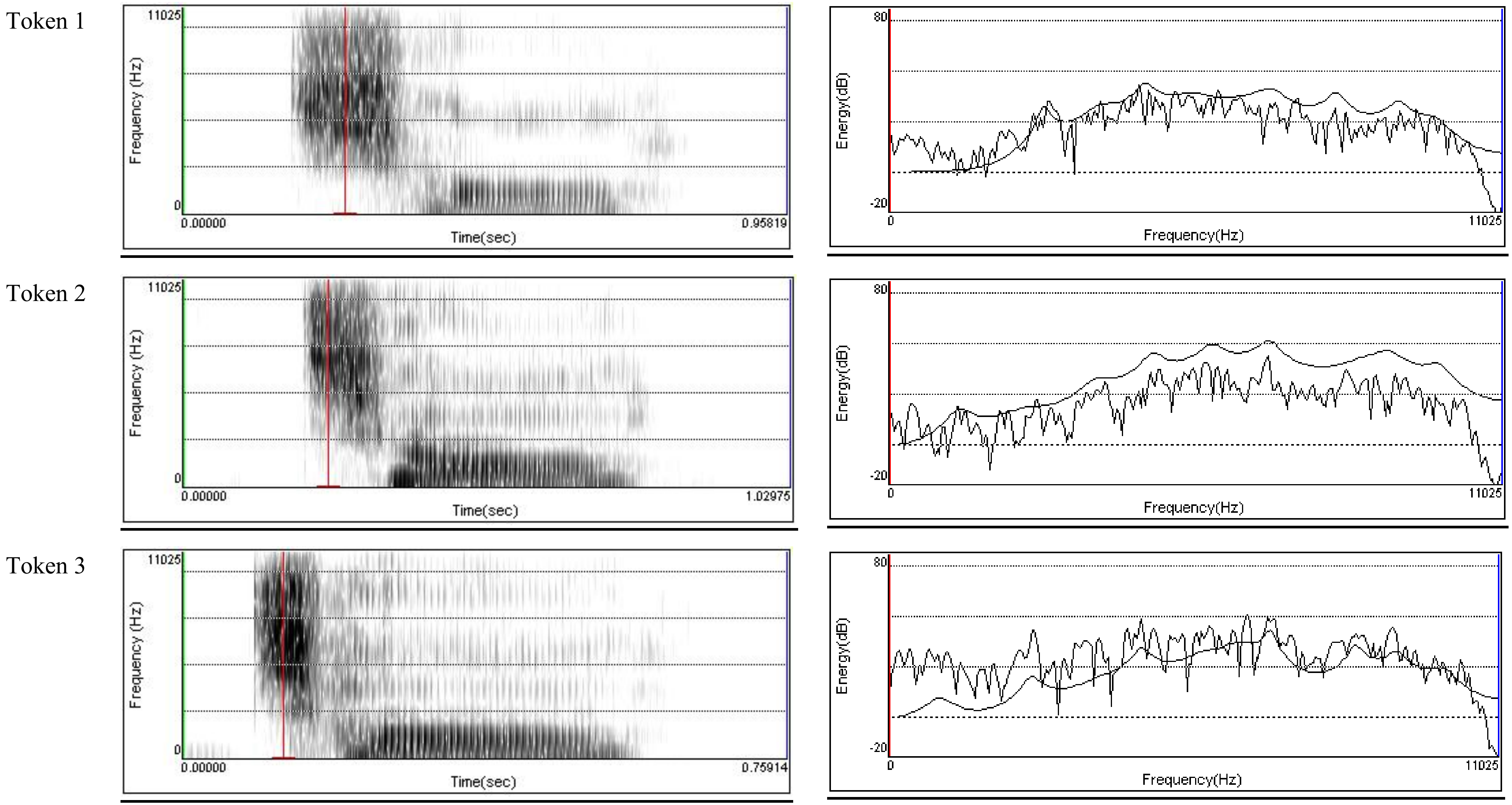


Figure 3.3.16. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [tɕʰ] in [tɕʰia˥] 掐 ‘nip off’

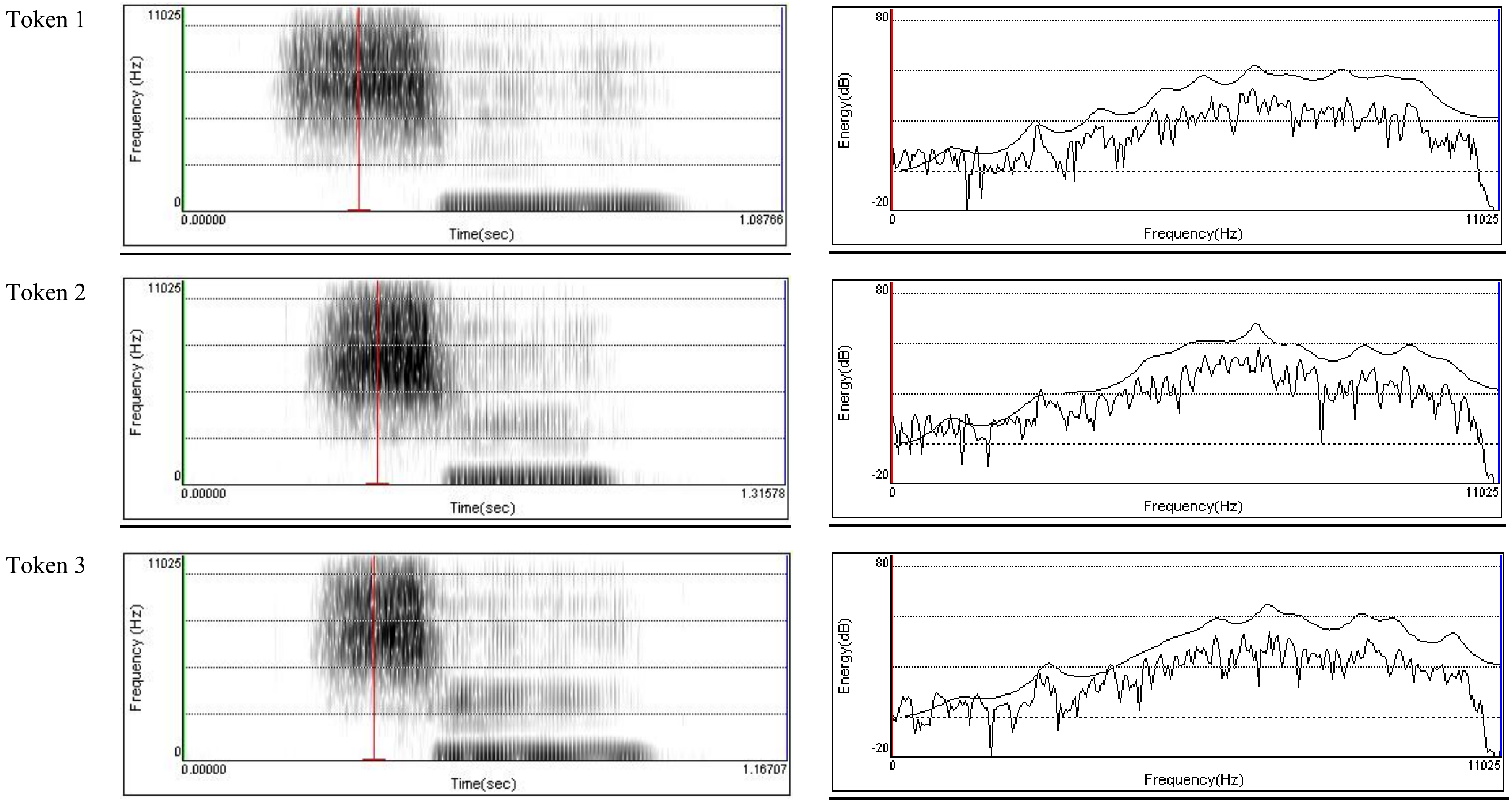


Figure 3.3.17. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [ɕ] in [ɕi˥] 希 ‘hope’

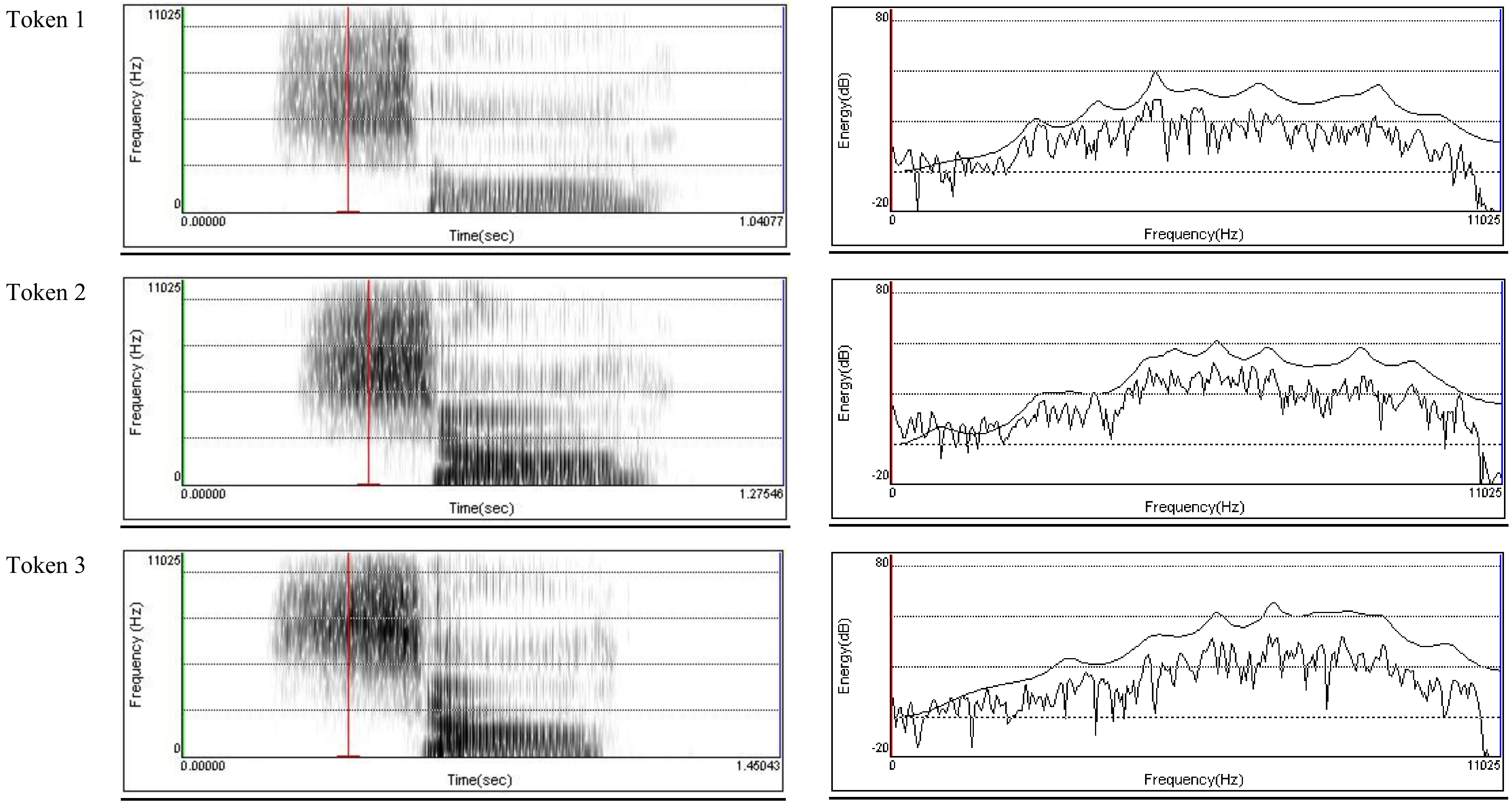


Figure 3.3.18. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [ɕ] in [ɕia˥] 蝦 'prawn'

3.4 Cantonese Female 1

Token 1

Token 2

Token 3

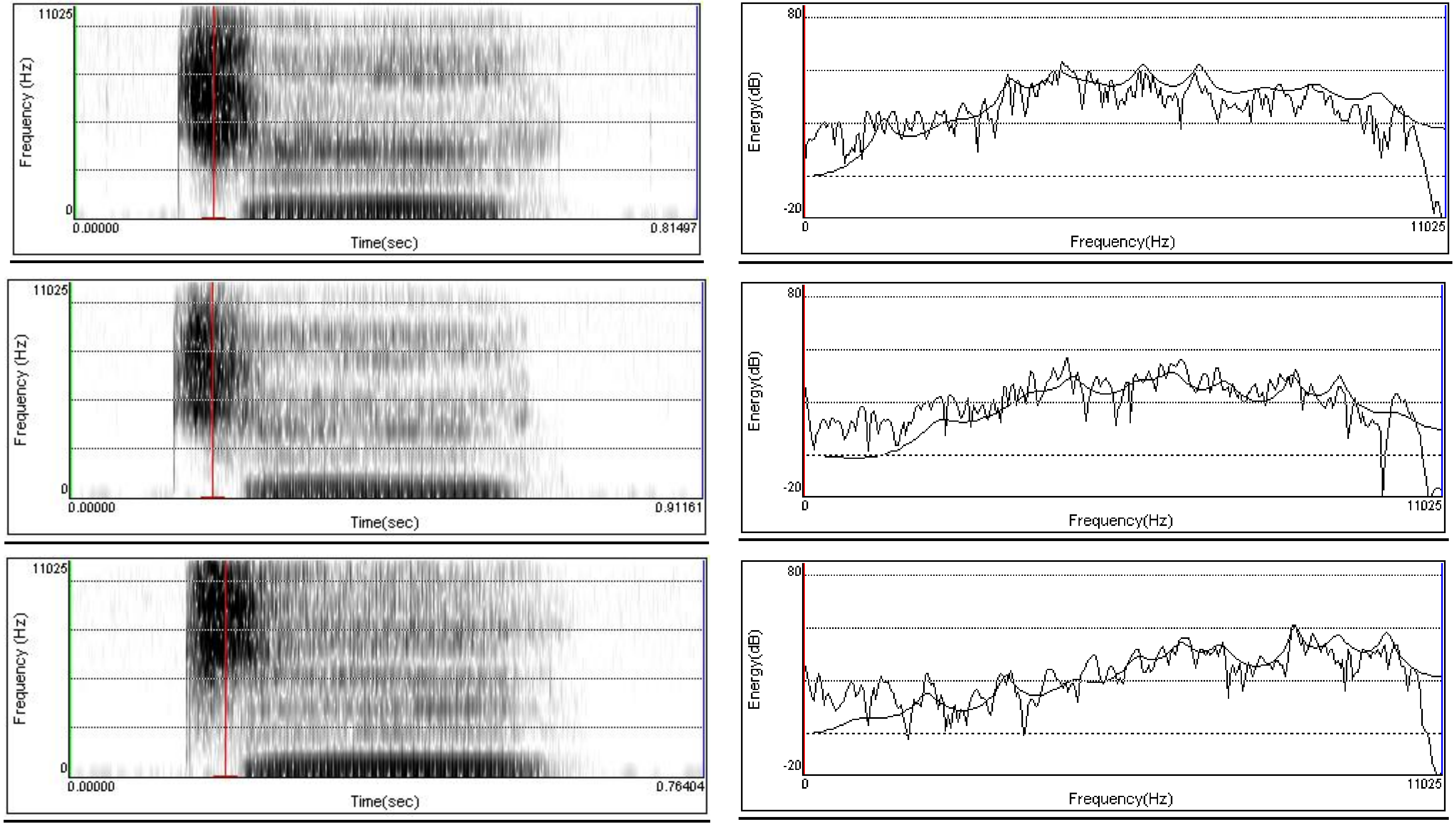


Figure 3.4.1. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [ʦ] in [ʦɿ˥] 資 ‘capital’

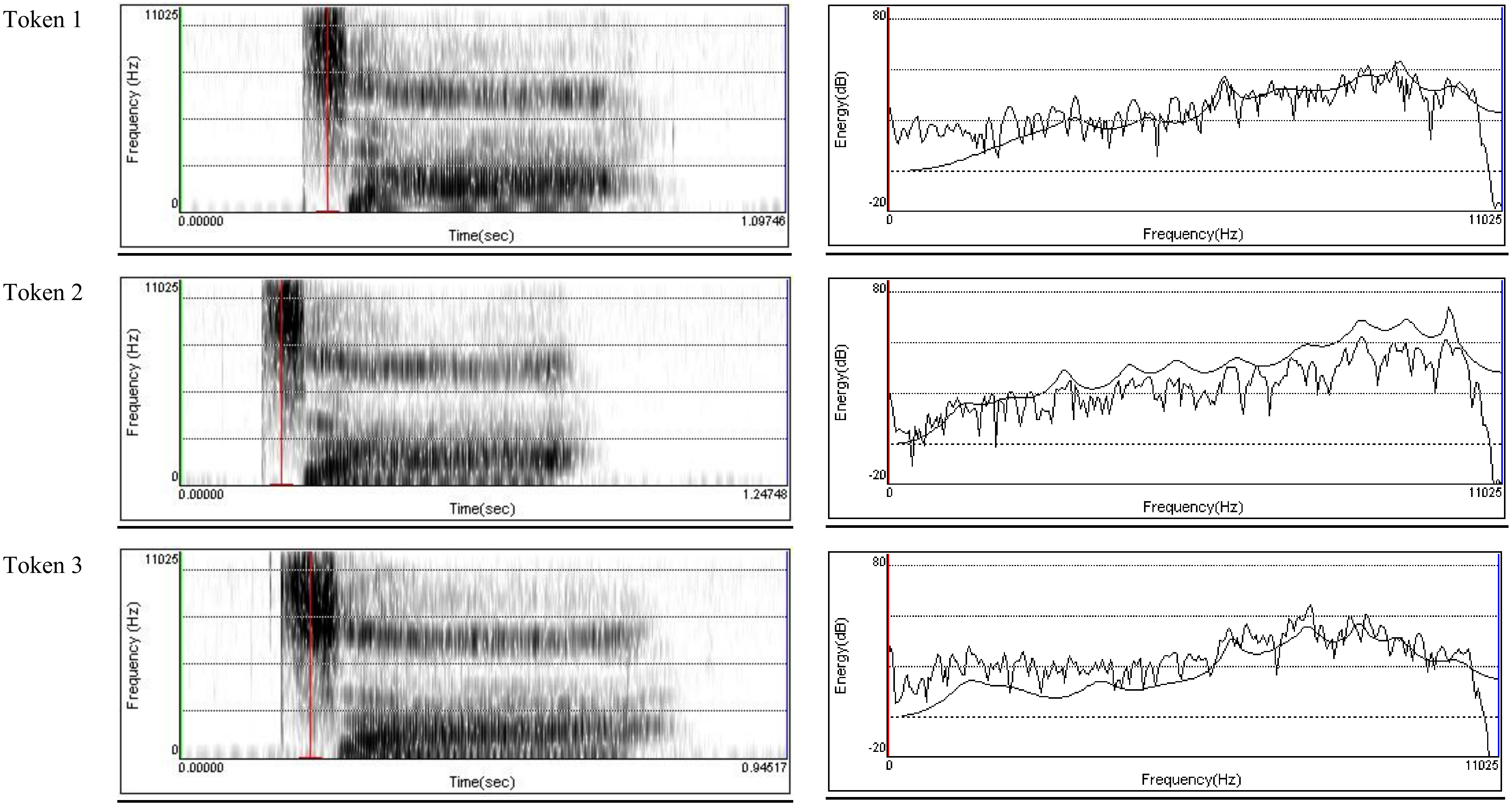


Figure 3.4.2. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [ʦ] in [ʦa˥] 紮 ‘tie’

Token 1

Token 2

Token 3

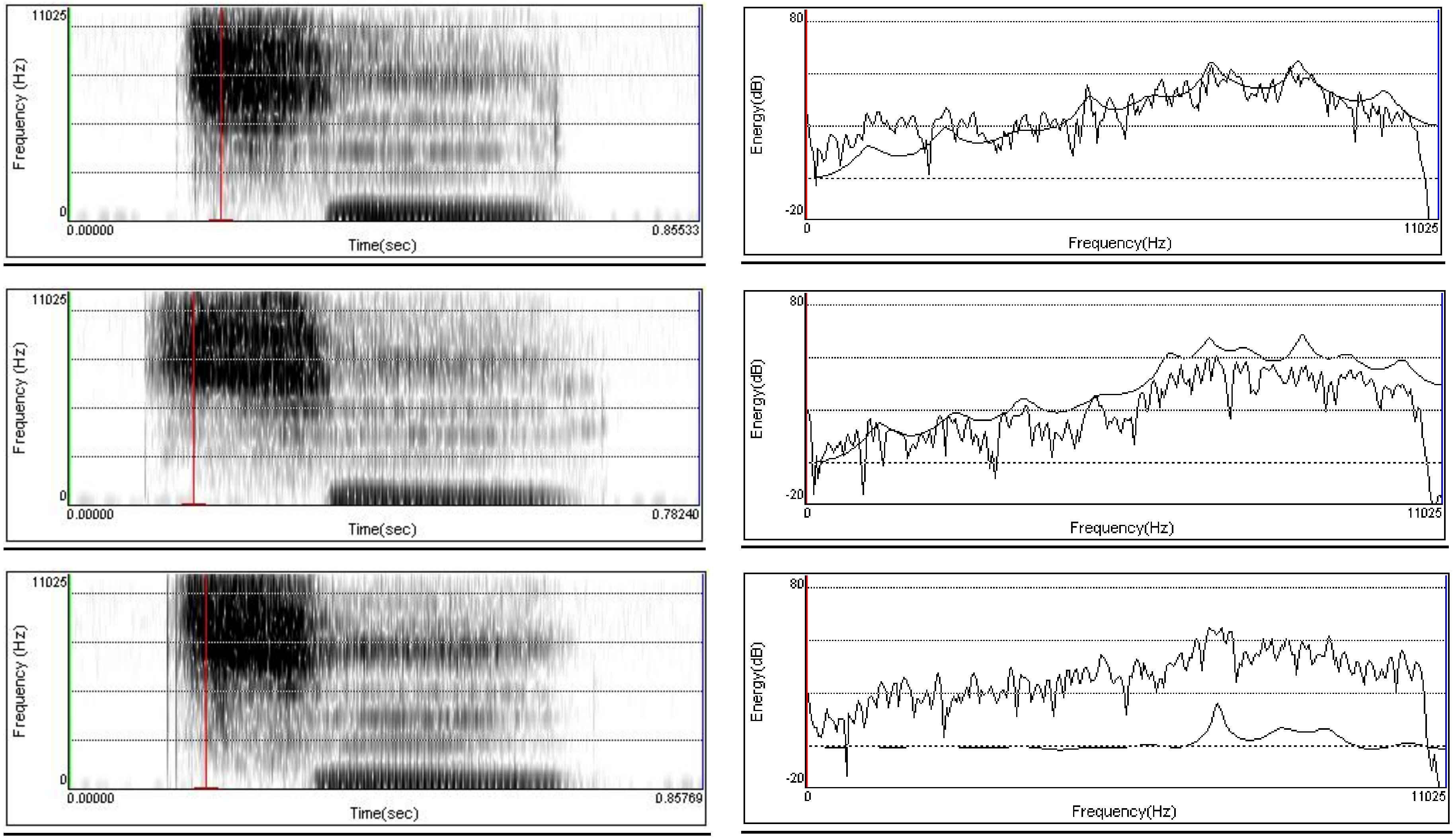


Figure 3.4.3. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [ʦʰ] in [ʦʰɿ˥] 疵 ‘flaw’

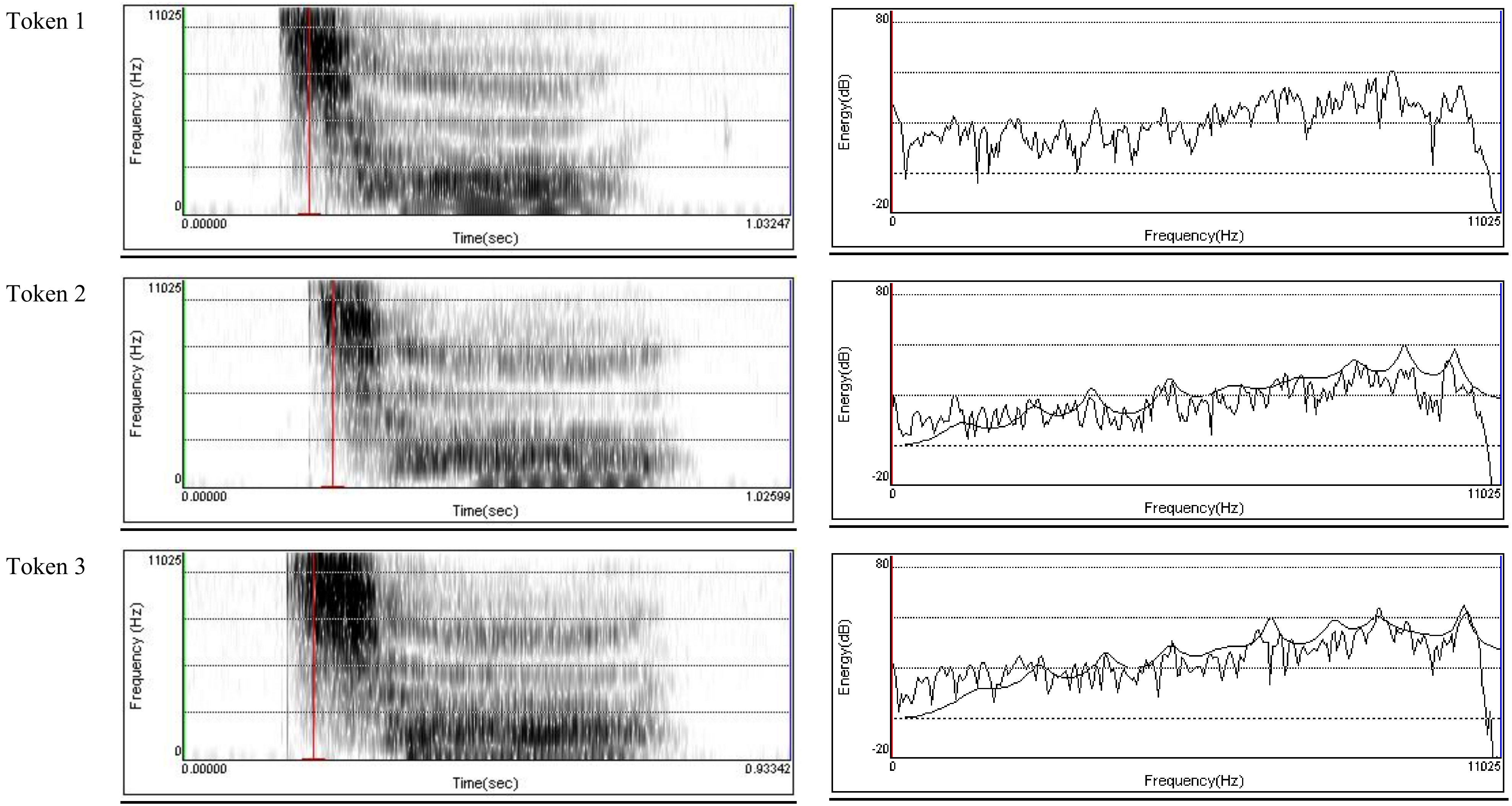


Figure 3.4.4. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [ʦʰ] in [ʦʰa˥] 擦 ‘rub’

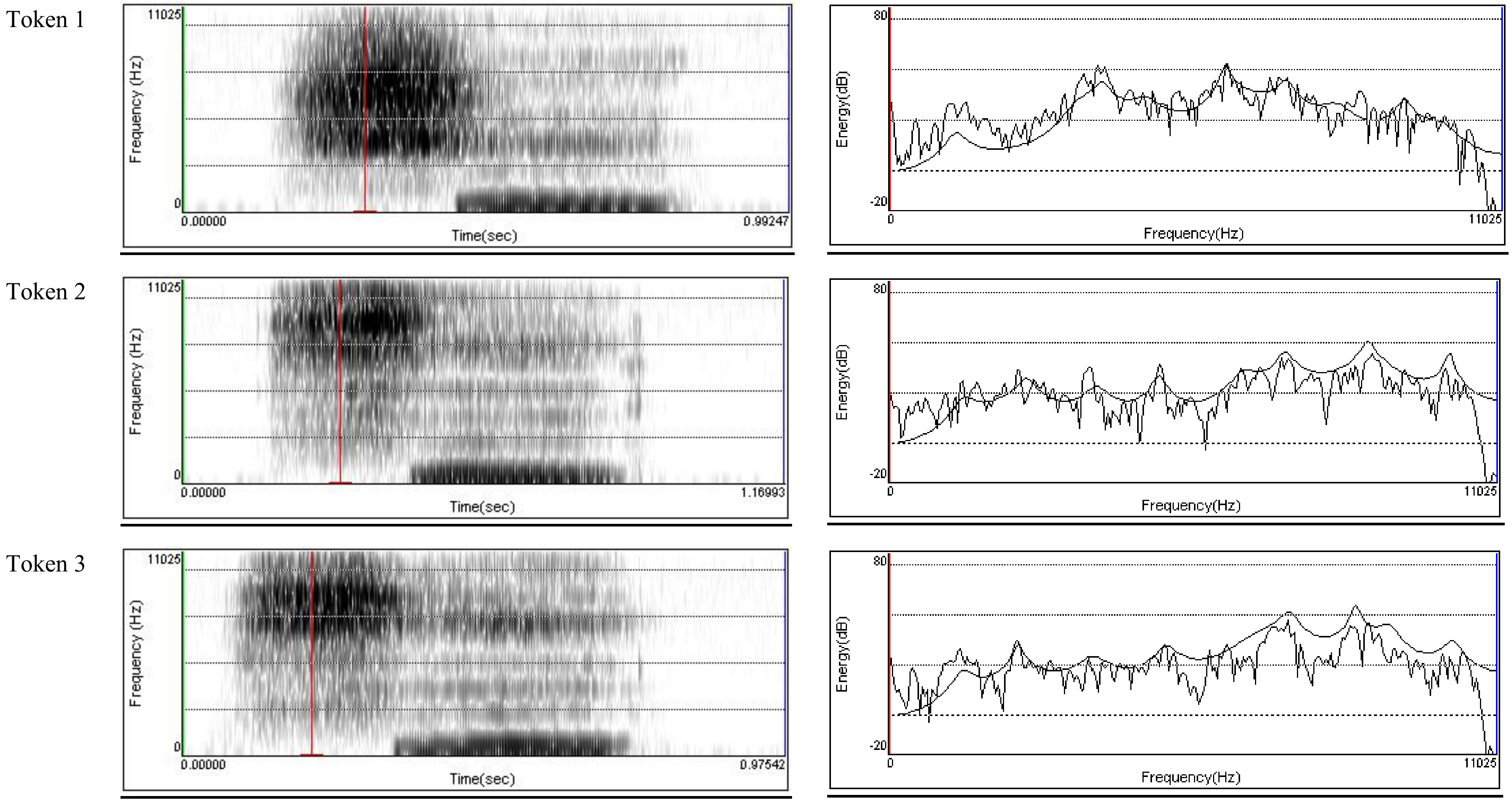


Figure 3.4.5. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [s] in [sɿ˥] 司 ‘in charge of’

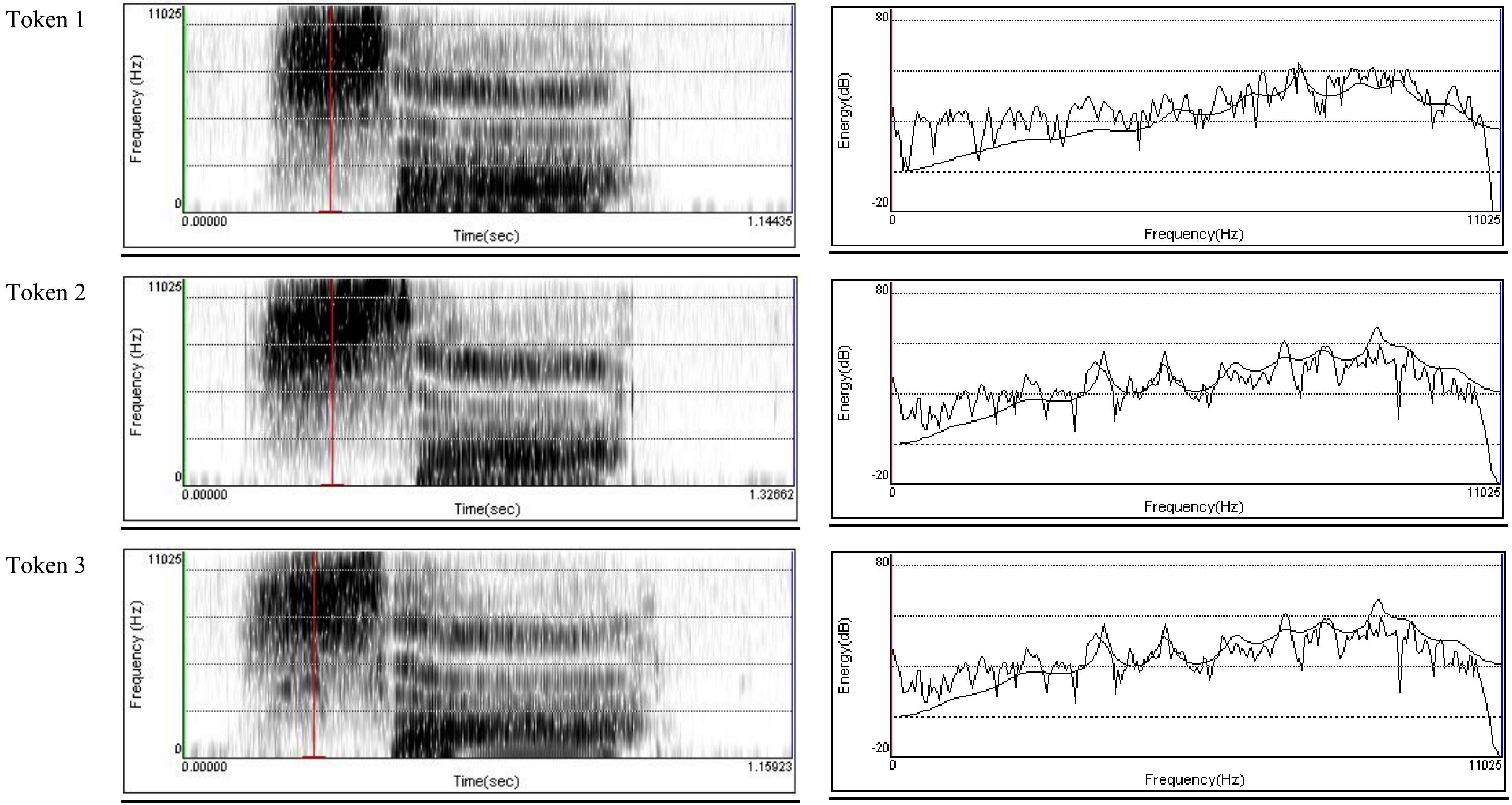


Figure 3.4.6. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [s] in [sa˥] 撒 ‘cast’

Token 1

Token 2

Token 3

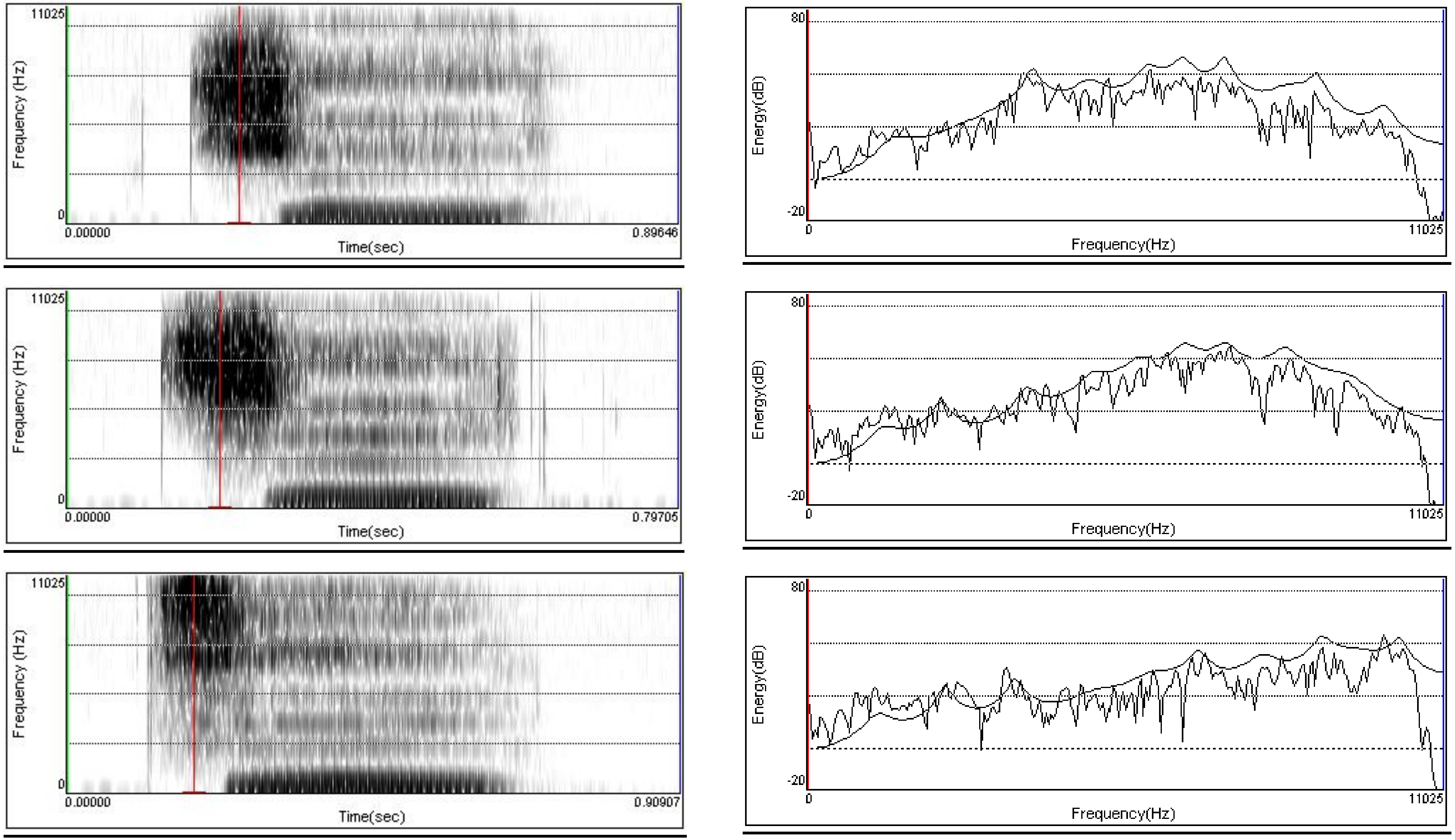


Figure 3.4.7. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [tʂ] in [tʂʅ˥] 知 'know'

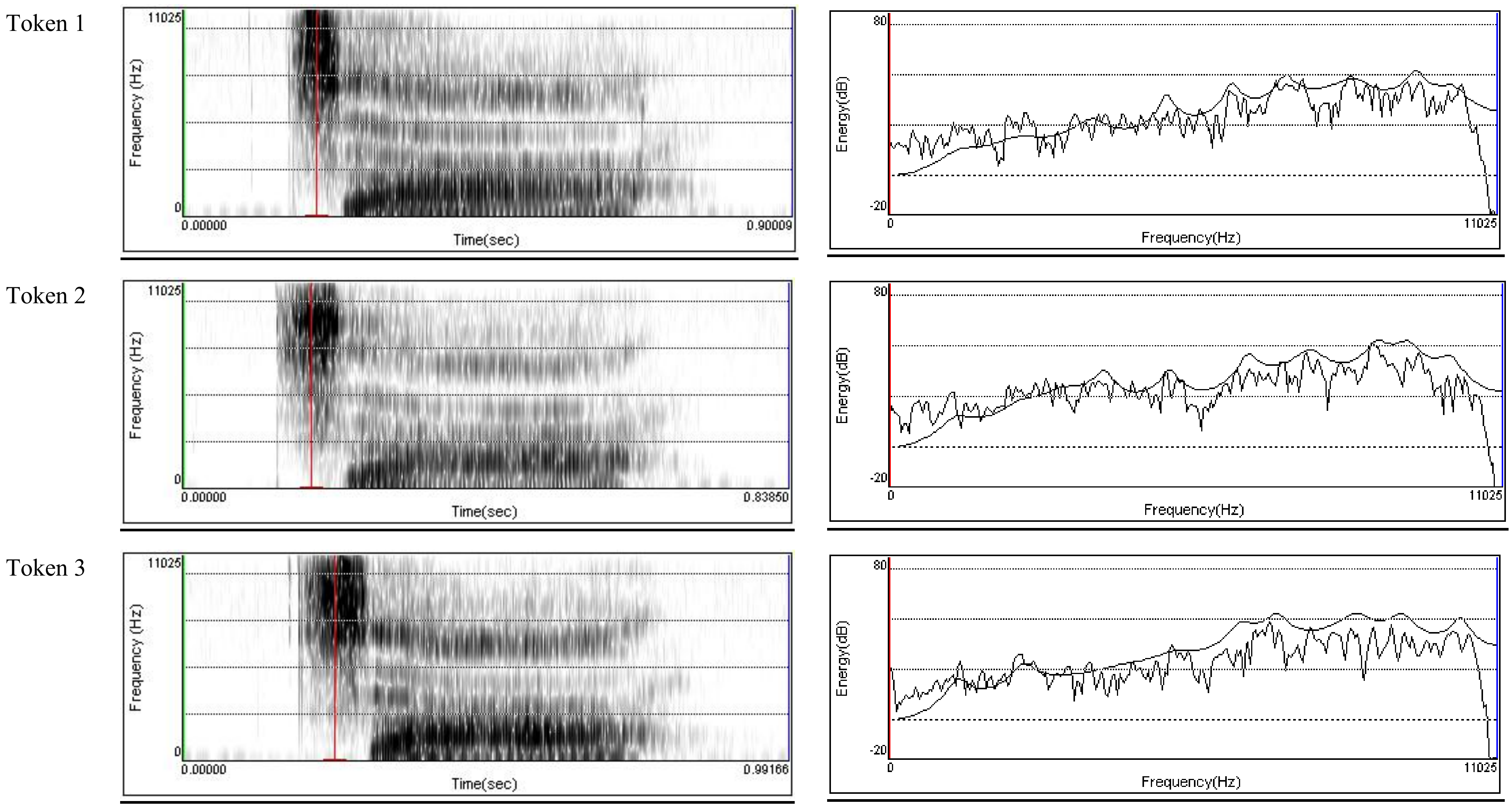


Figure 3.4.8. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [tʂ] in [tʂa˥] 渣 ‘residue’

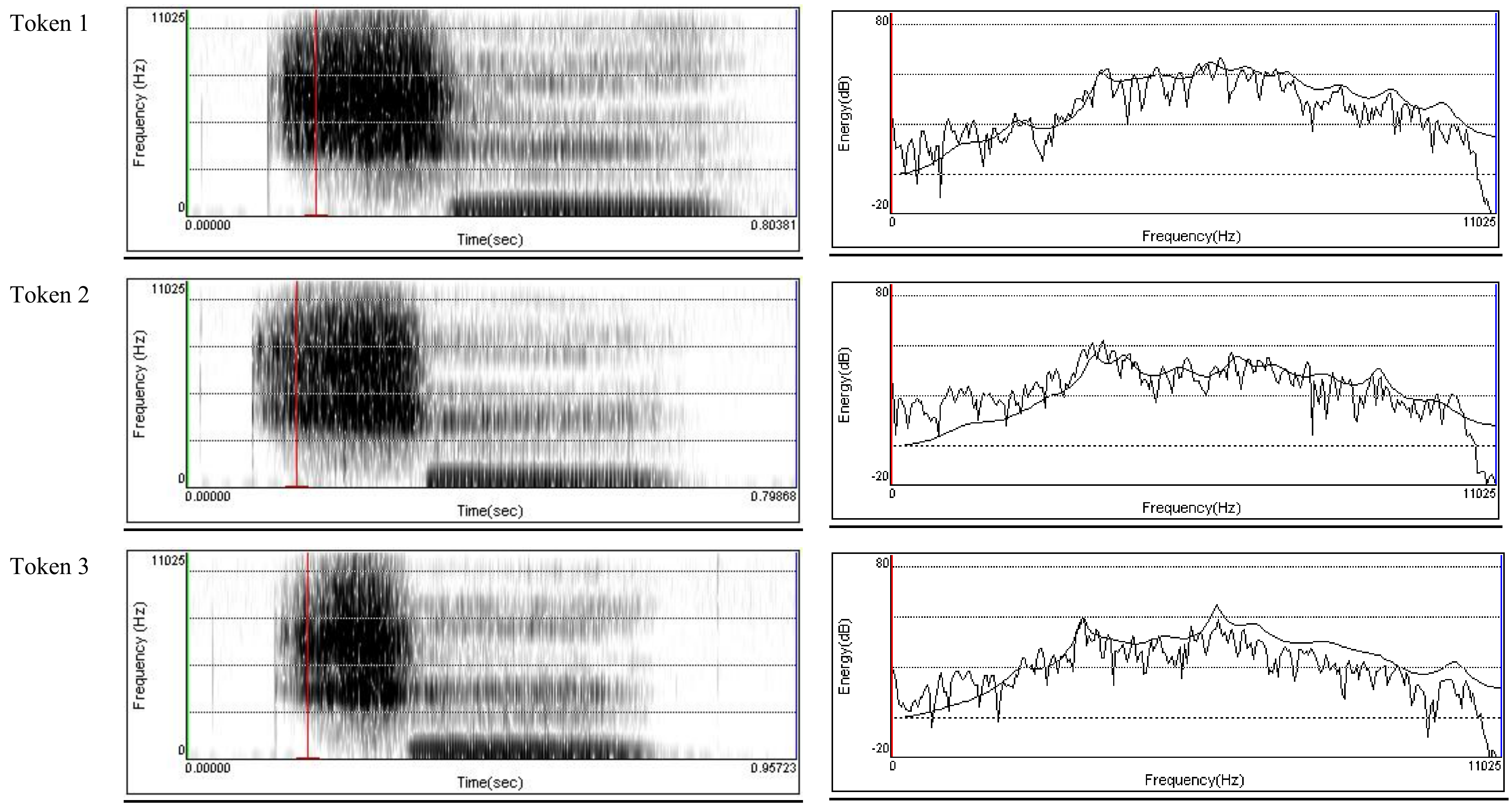


Figure 3.4.9. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [tʂʰ] in [tʂʰʅ˥] 吃 ‘eat’

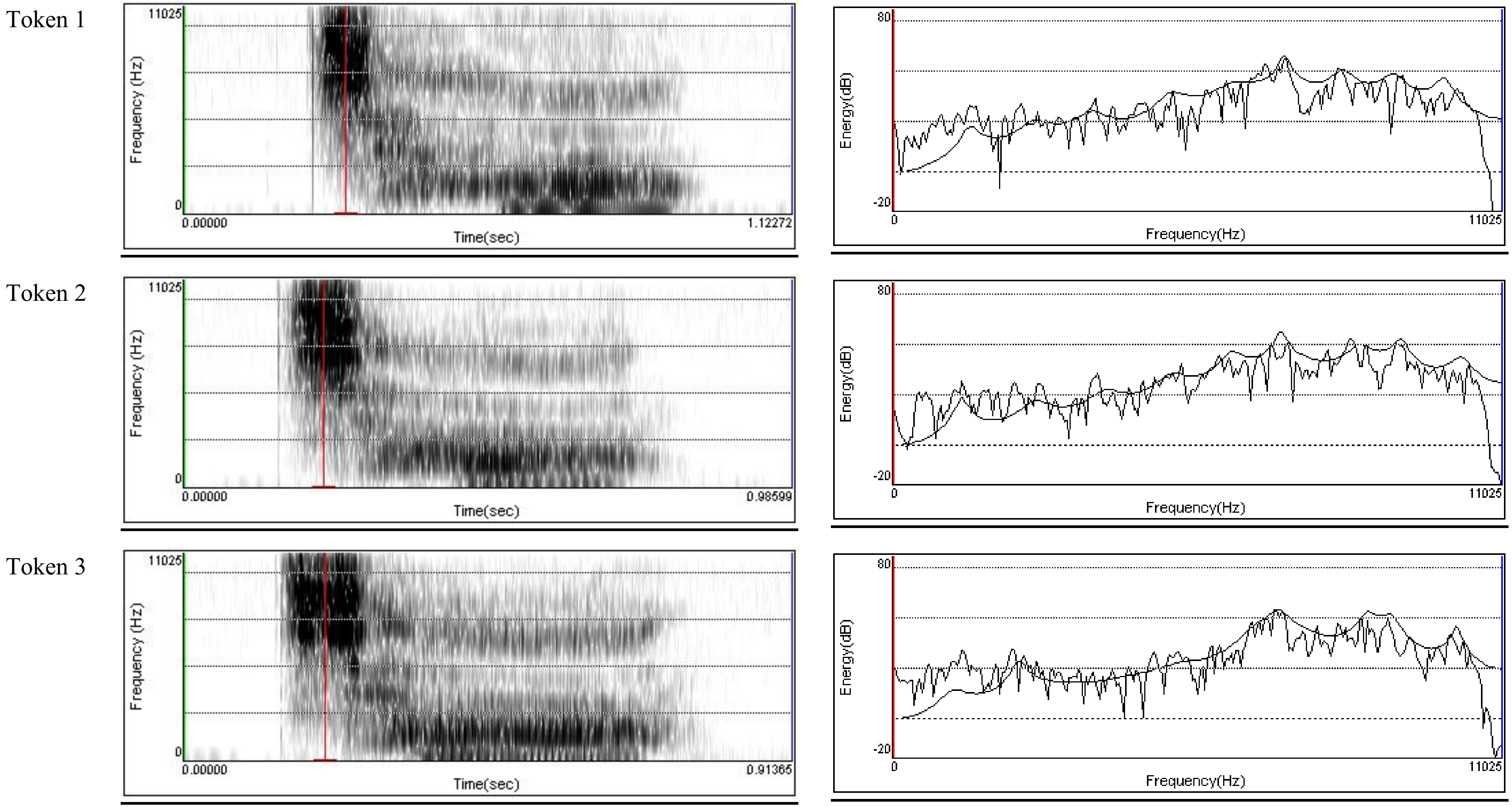


Figure 3.4.10. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [tʂʰ] in [tʂʰa˥] 叉 ‘folk’

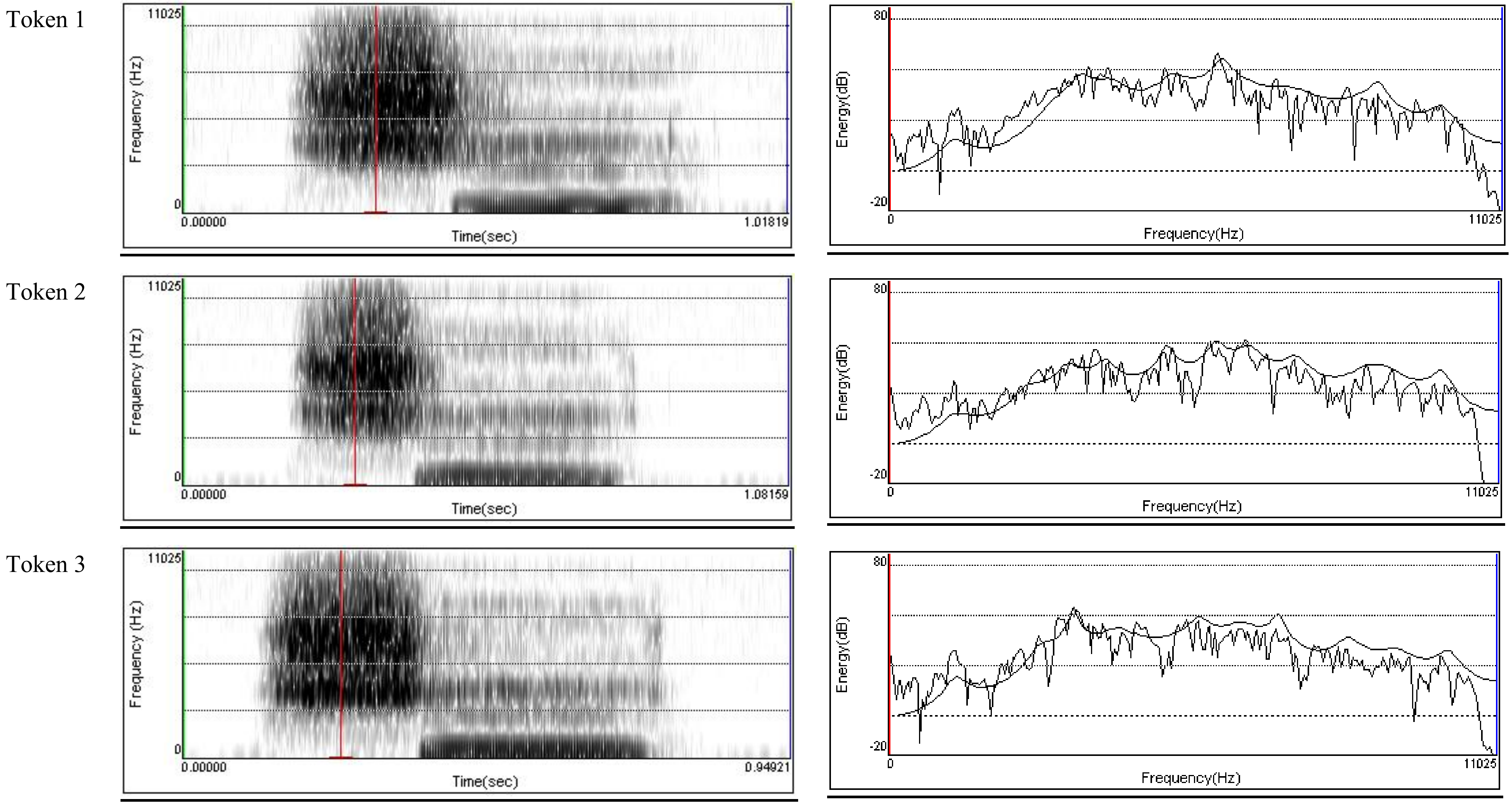


Figure 3.4.11. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [ʂ] in [ʂʅ˥] 失 ‘lose’

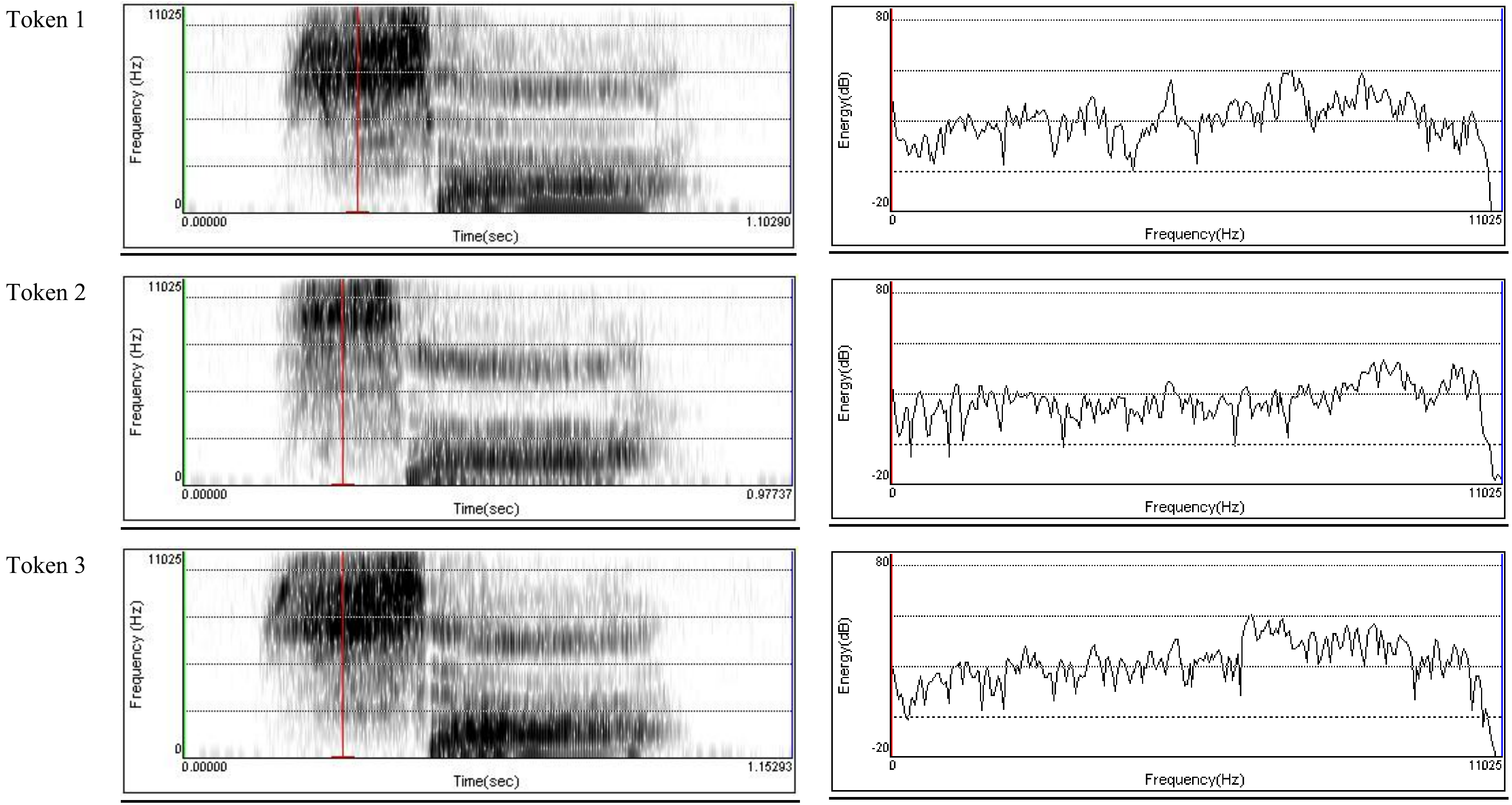


Figure 3.4.12. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [ʂ] in [ʂa˥] 沙 ‘sand’

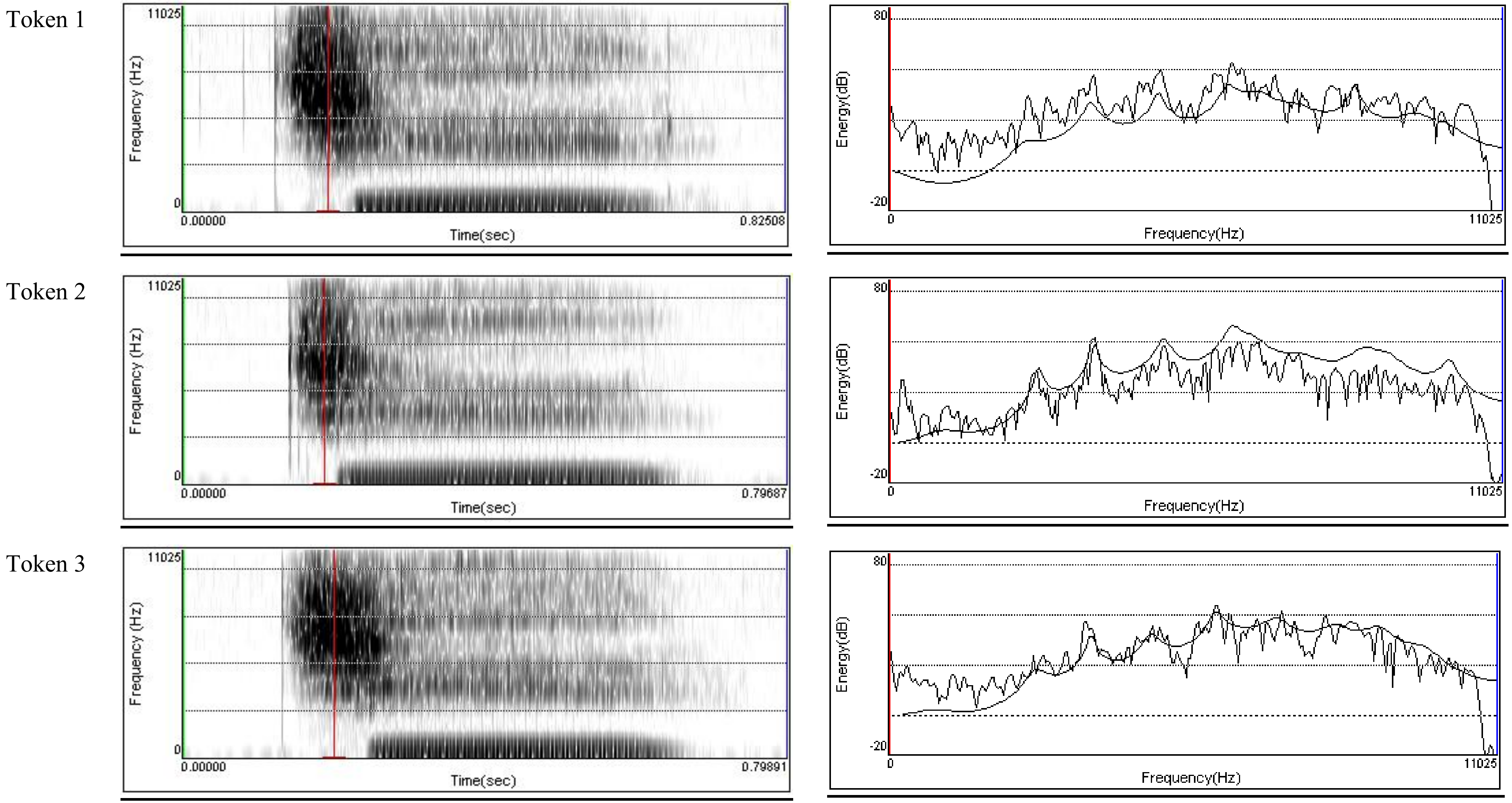


Figure 3.4.13. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [tɕ] in [tɕi˥] 基 ‘base’

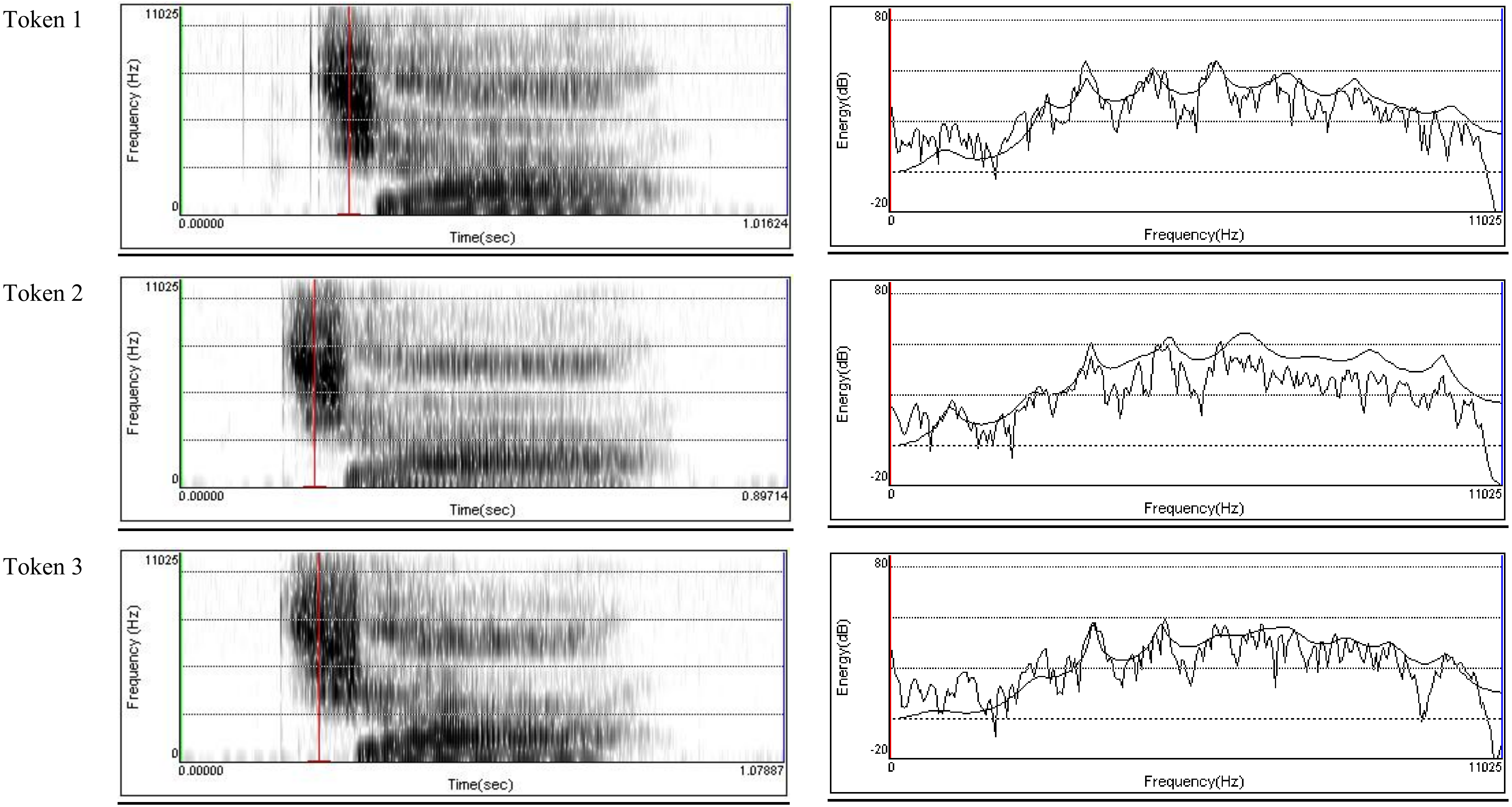


Figure 3.4.14. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [tɕ] in [tɕia˥] 加 'add'

Token 1

Token 2

Token 3

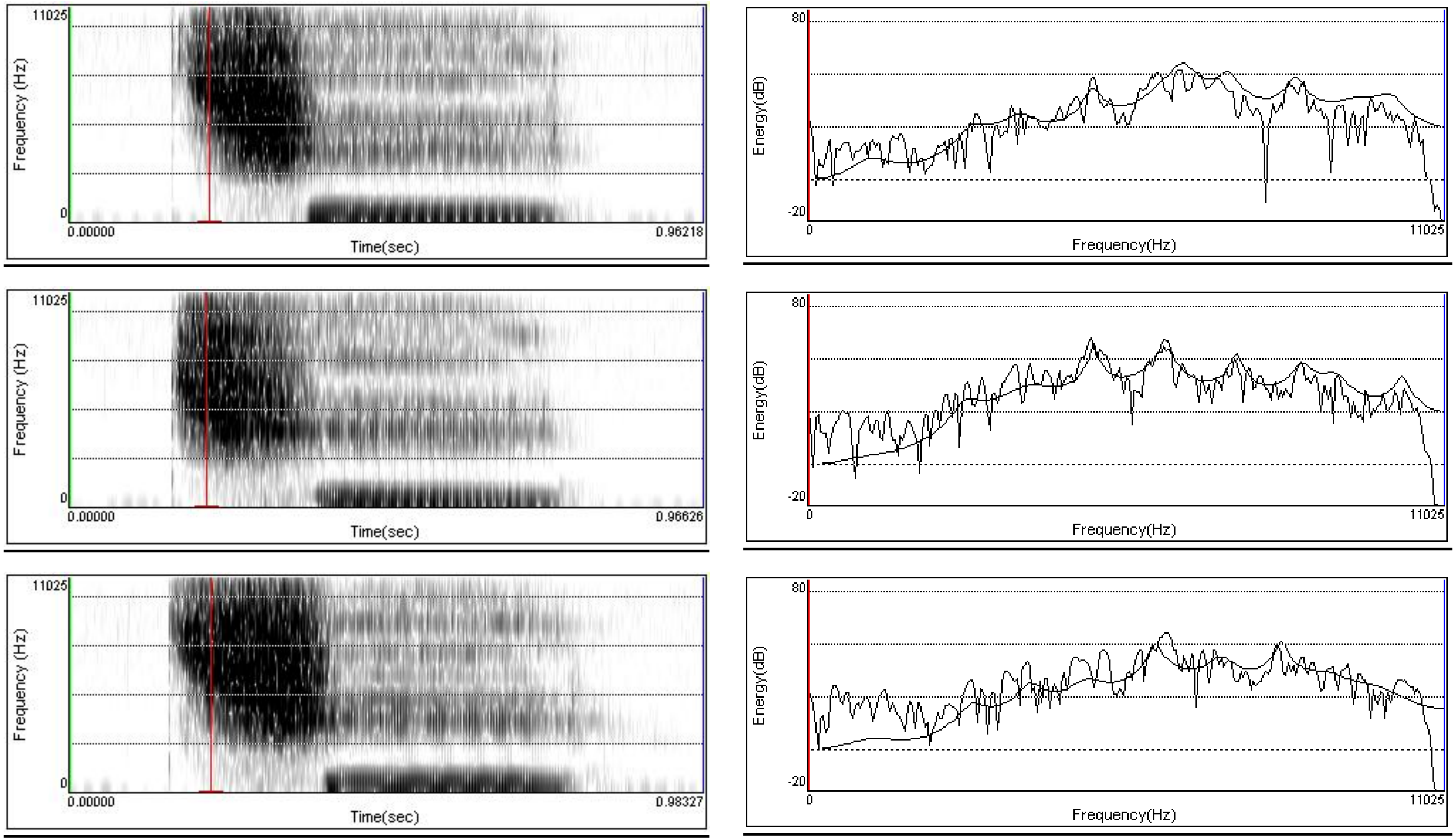


Figure 3.4.15. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [tɕʰ] in [tɕʰi˥] 七 'seven'

Token 1

Token 2

Token 3

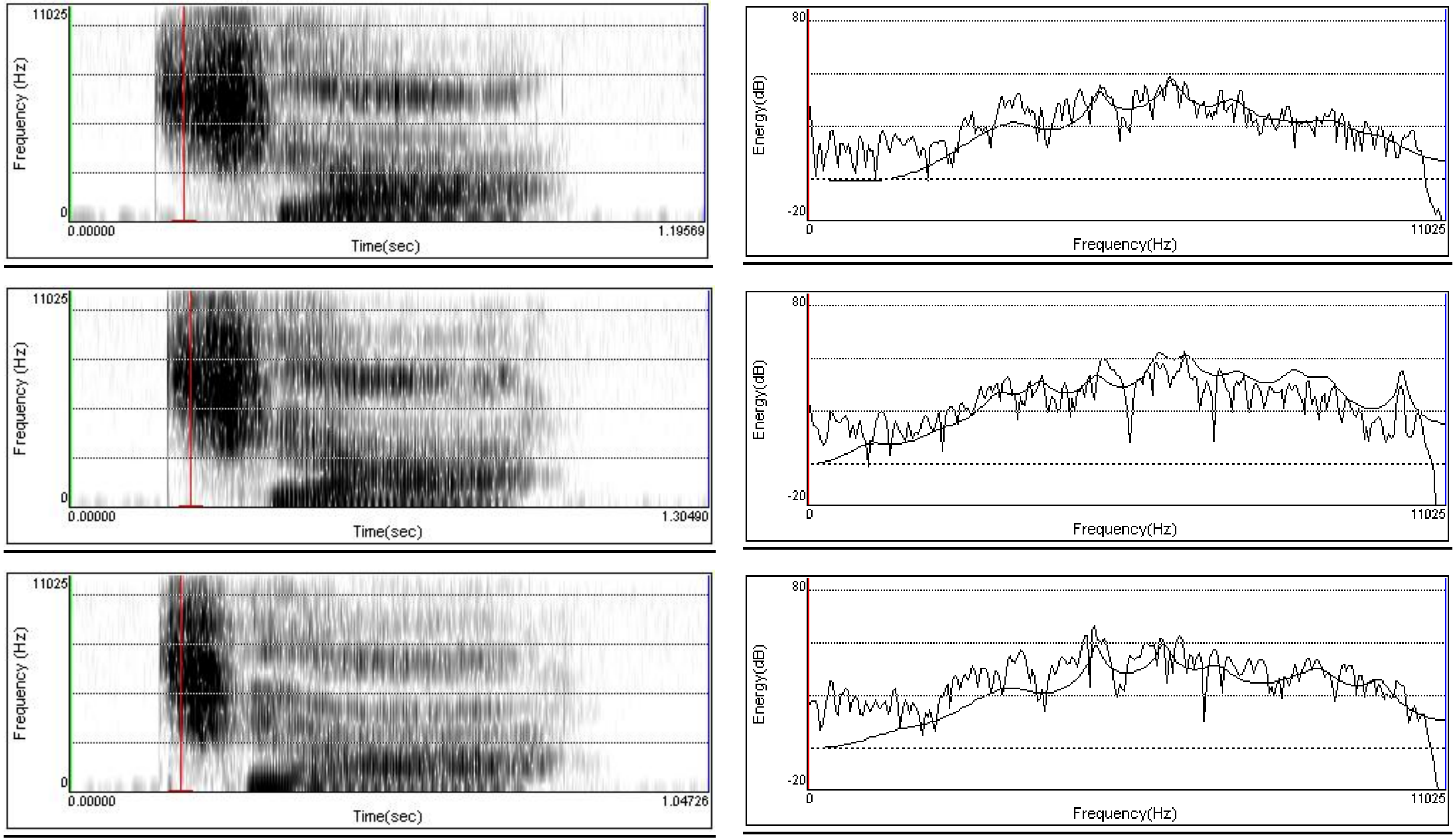


Figure 3.4.16. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [tɕʰ] in [tɕʰia˥] 掐 ‘nip off’

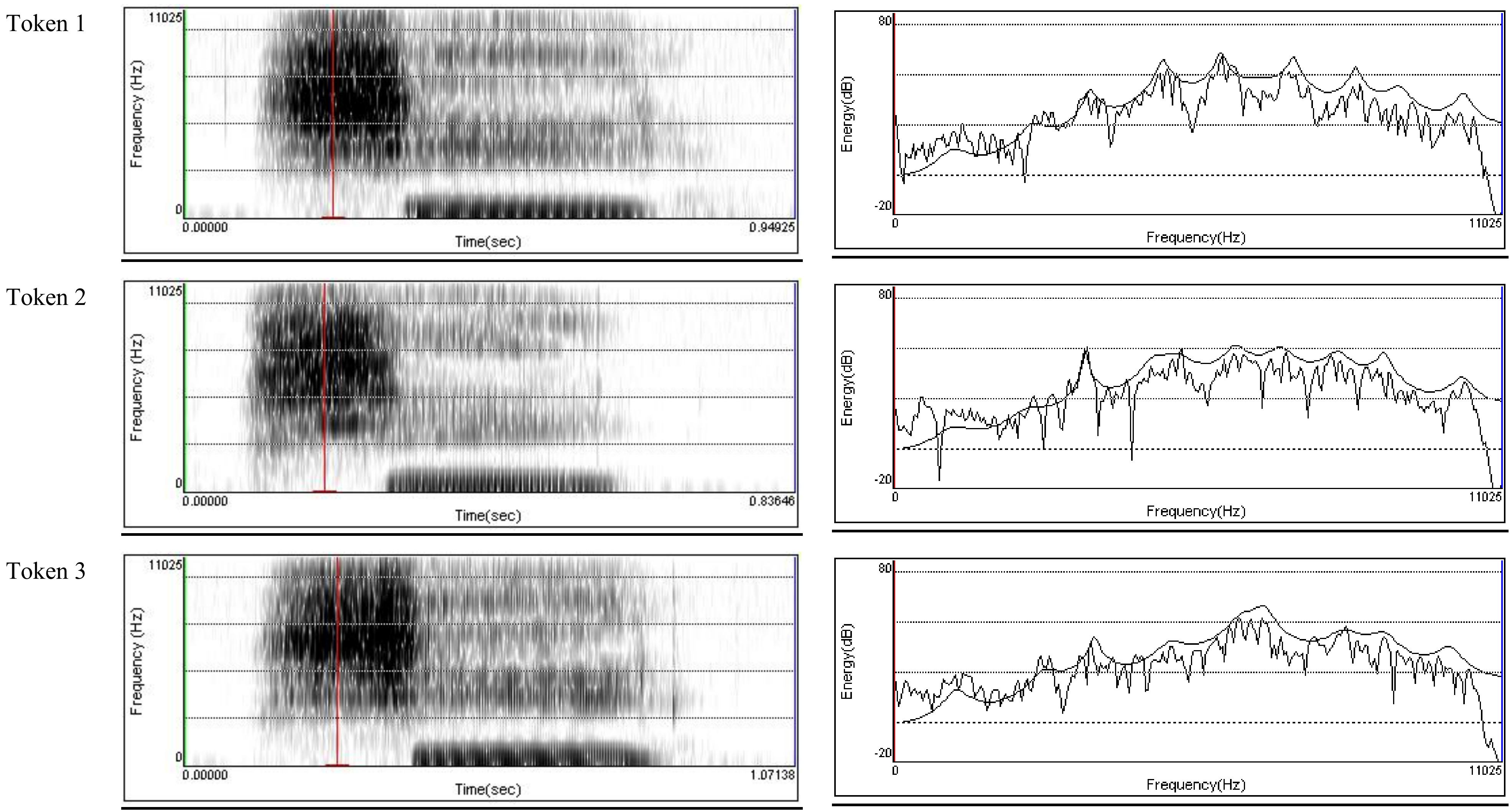


Figure 3.4.17. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [ɕ] in [ɕi˥] 希 ‘hope’

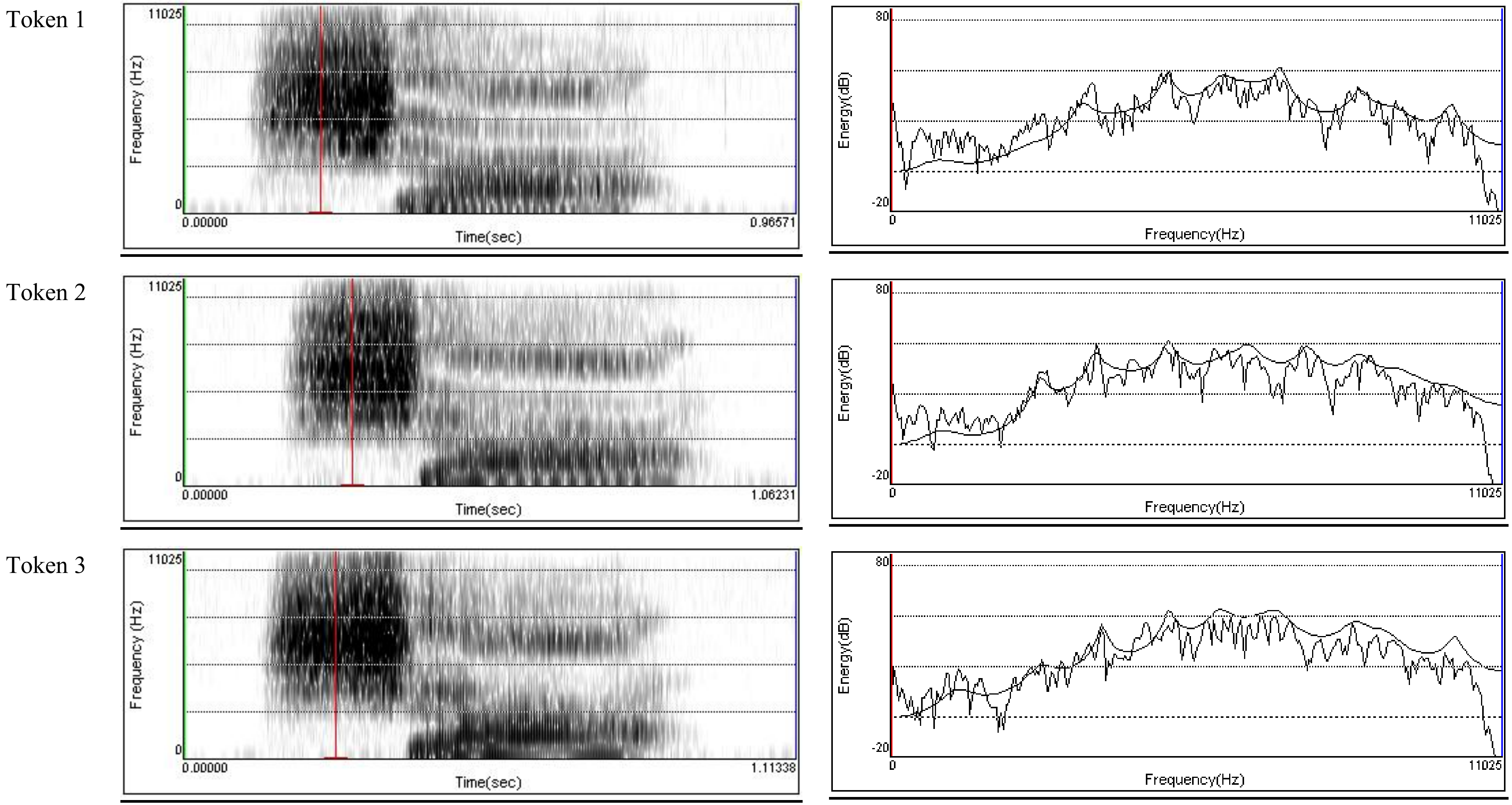


Figure 3.4.18. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [ɕ] in [ɕia˥] 蝦 'prawn'

3.5 Cantonese Female 2

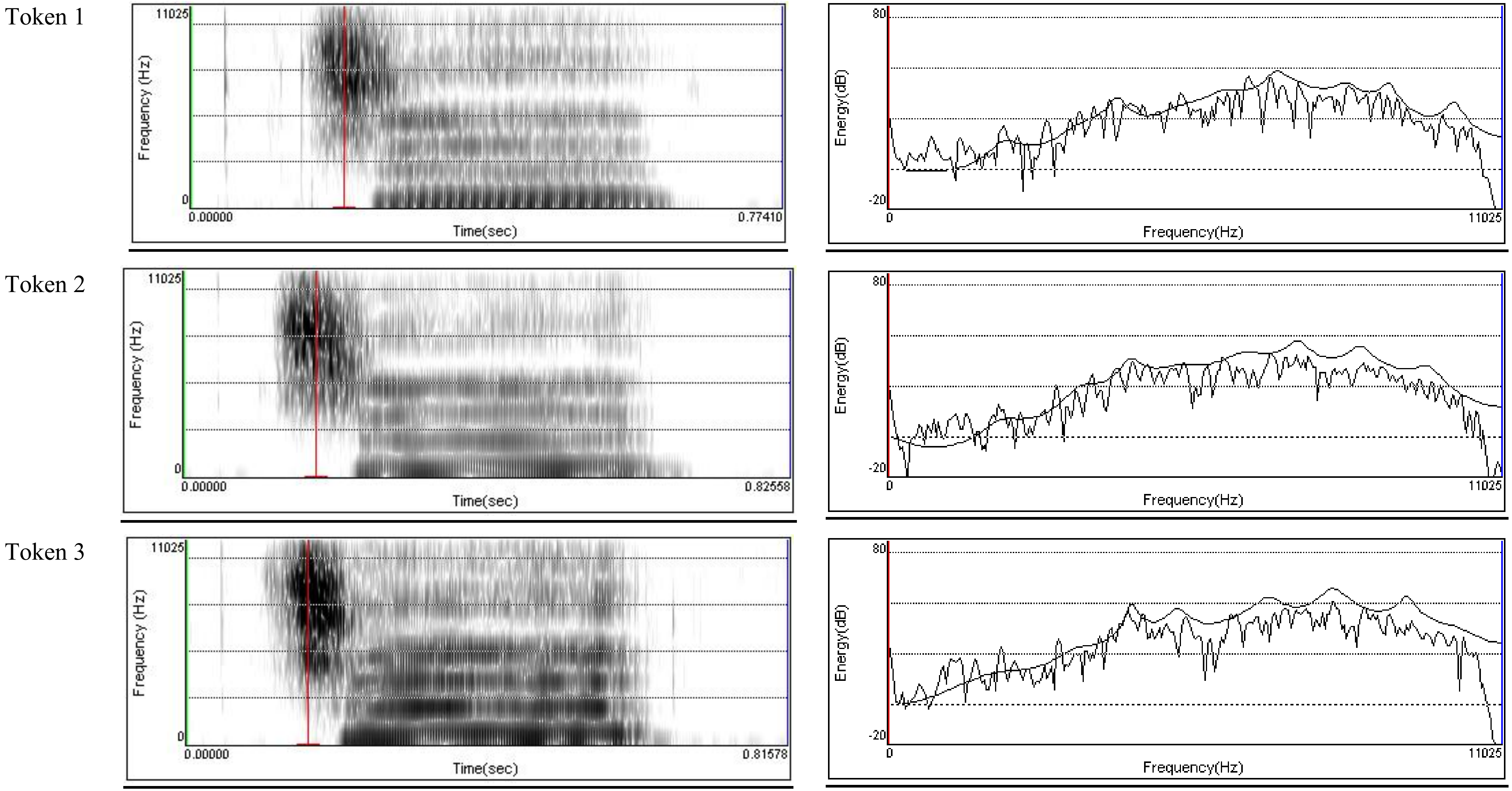


Figure 3.5.1. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [ʦ] in [ʦɿ˥] 資 'capital'

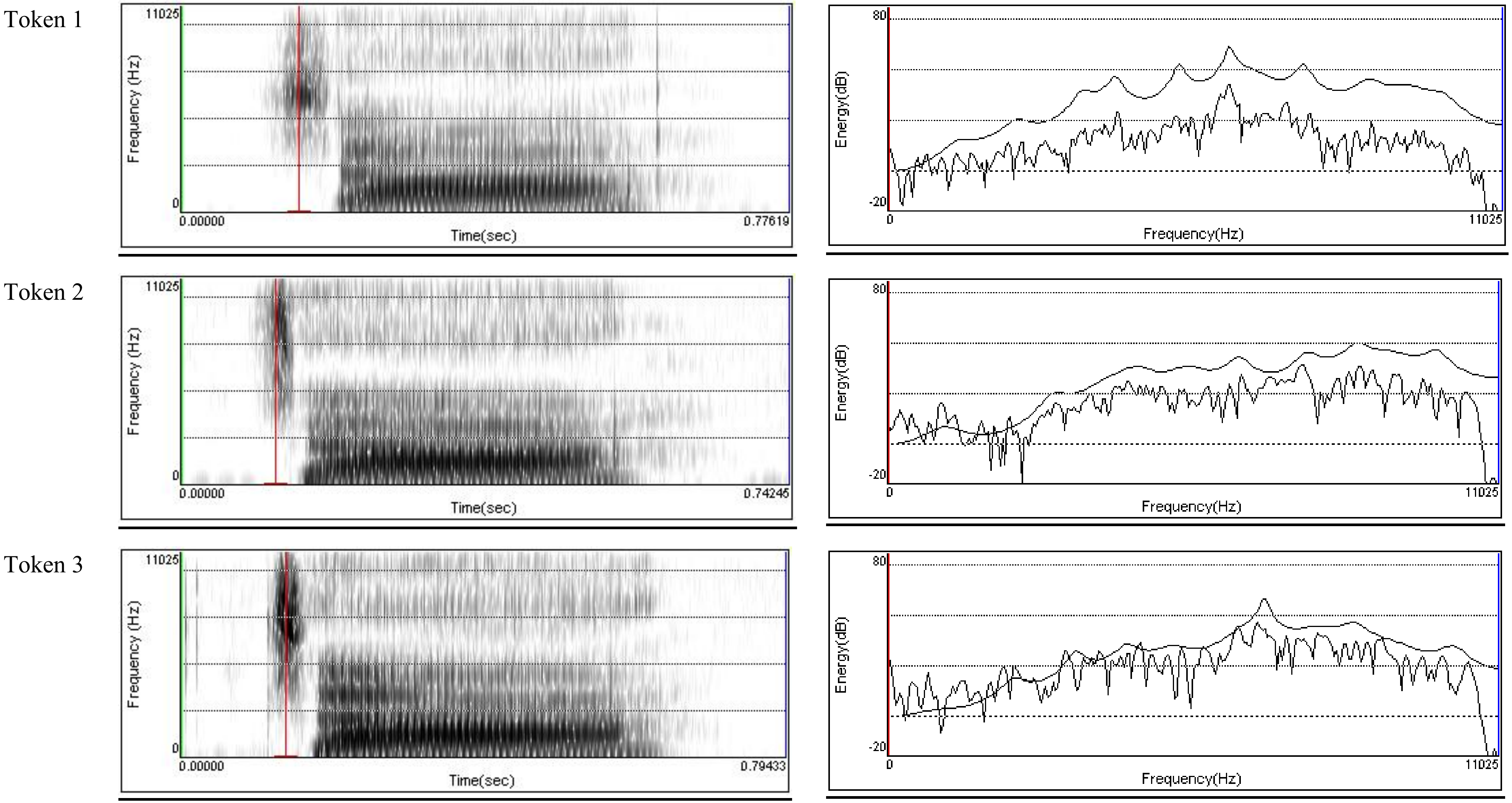


Figure 3.5.2. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [ʦ] in [ʦa˥] 紮 'tie'

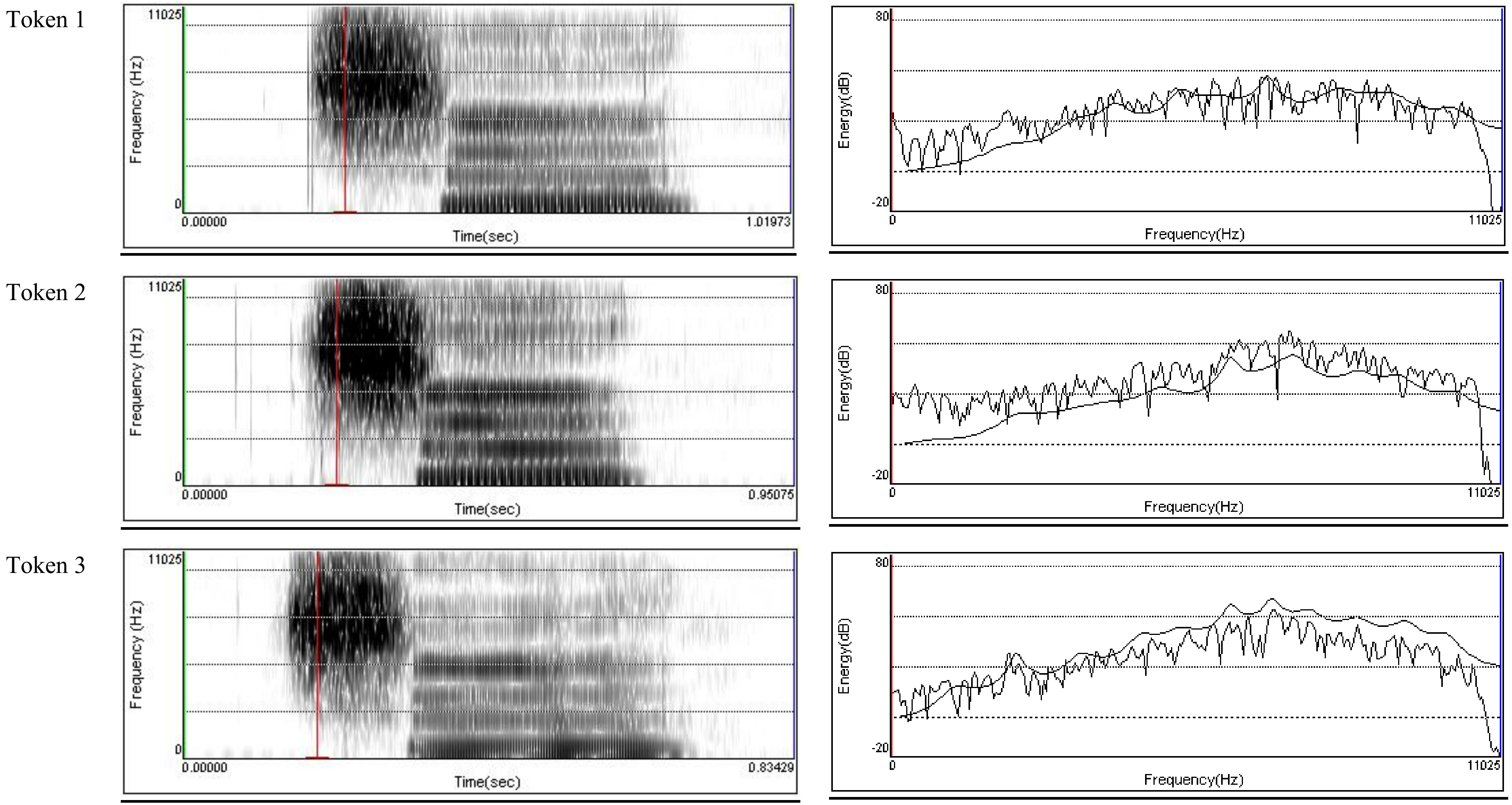
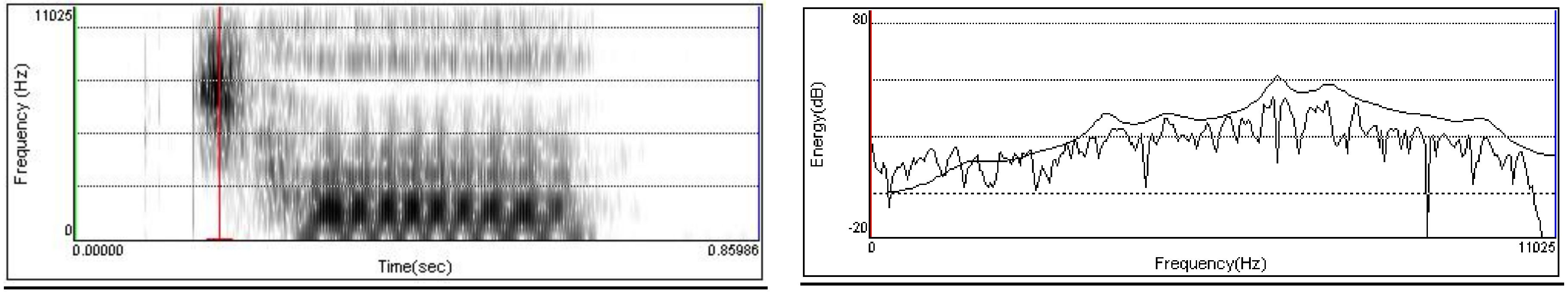
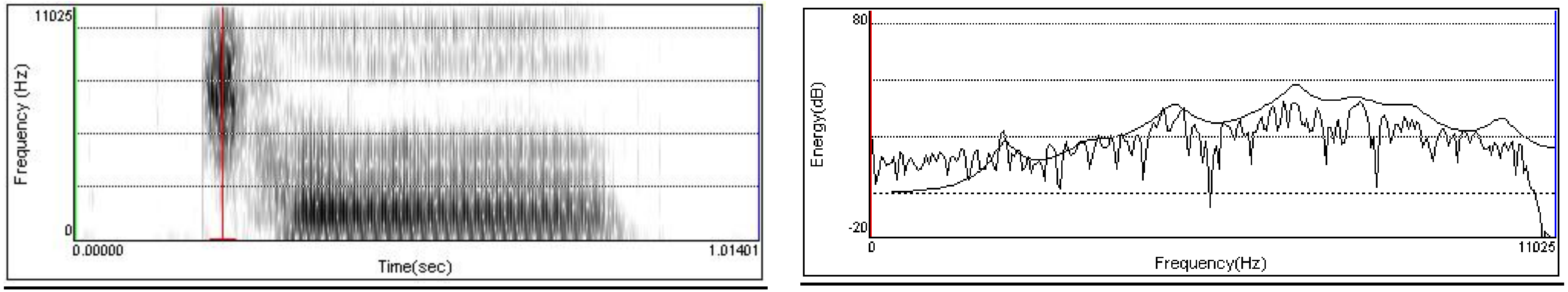


Figure 3.5.3. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [ʦʰ] in [ʦʰɿ˥] 疵 'flaw'

Token 1



Token 2



Token 3

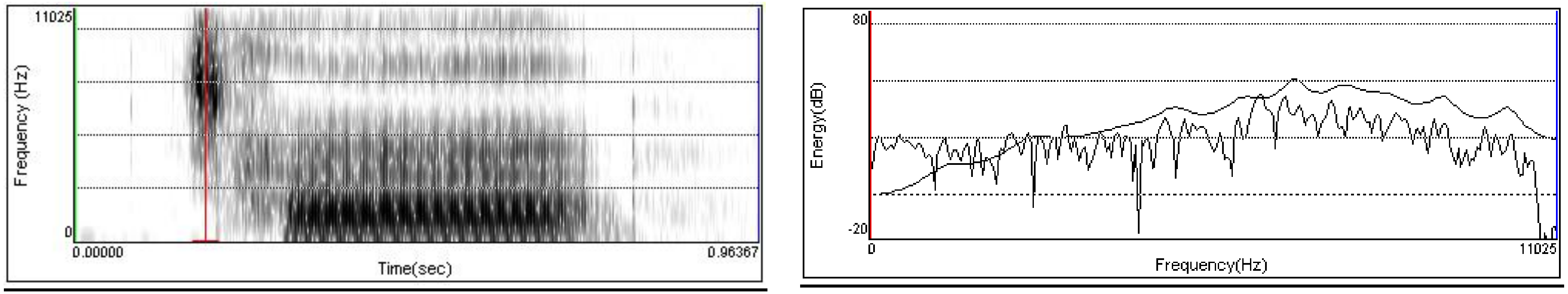


Figure 3.5.4. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [ʦʰ] in [ʦʰa˥] 擦 ‘rub’

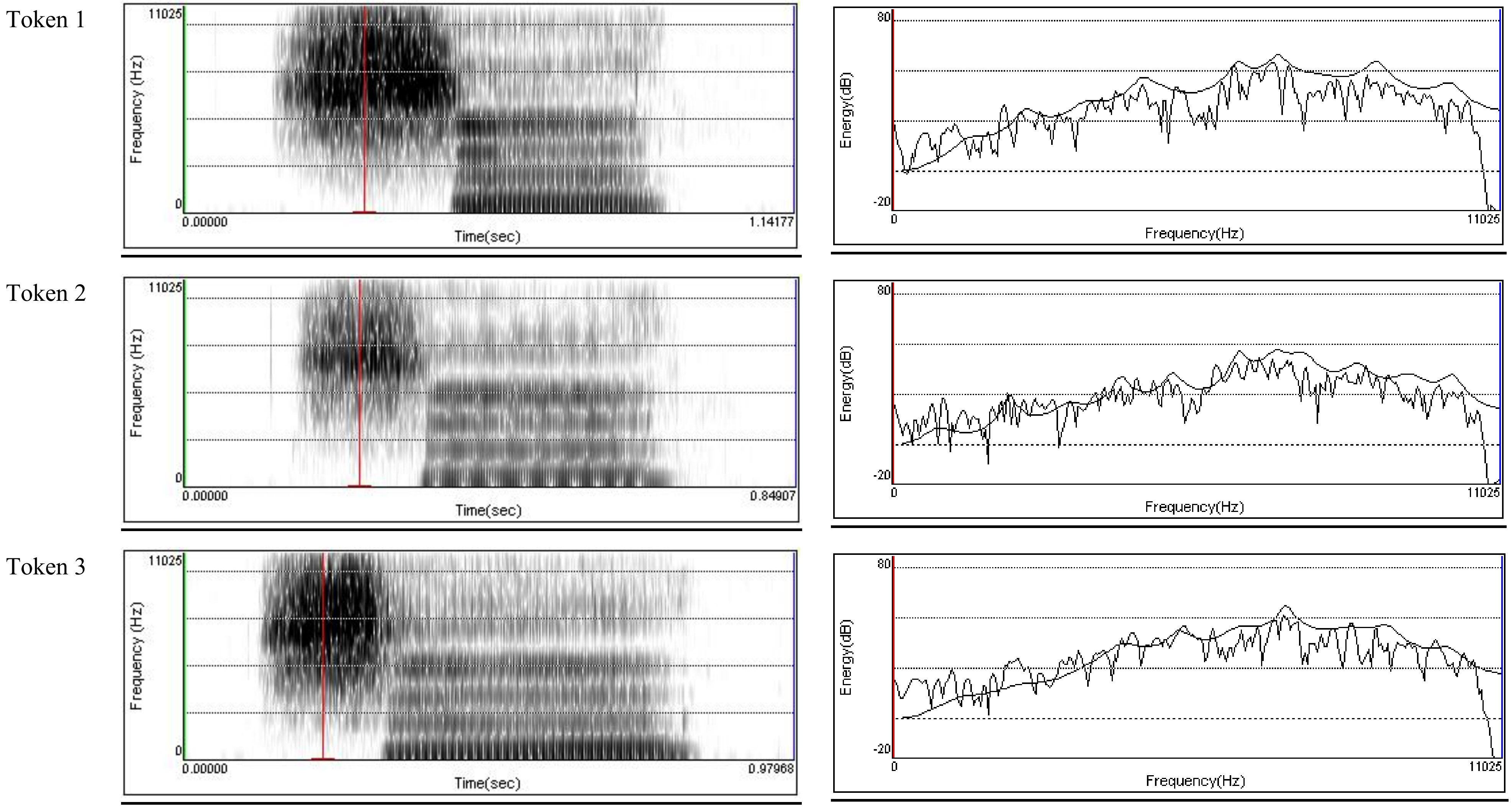


Figure 3.5.5. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [s] in [sɿ˥] 司 ‘in charge of’

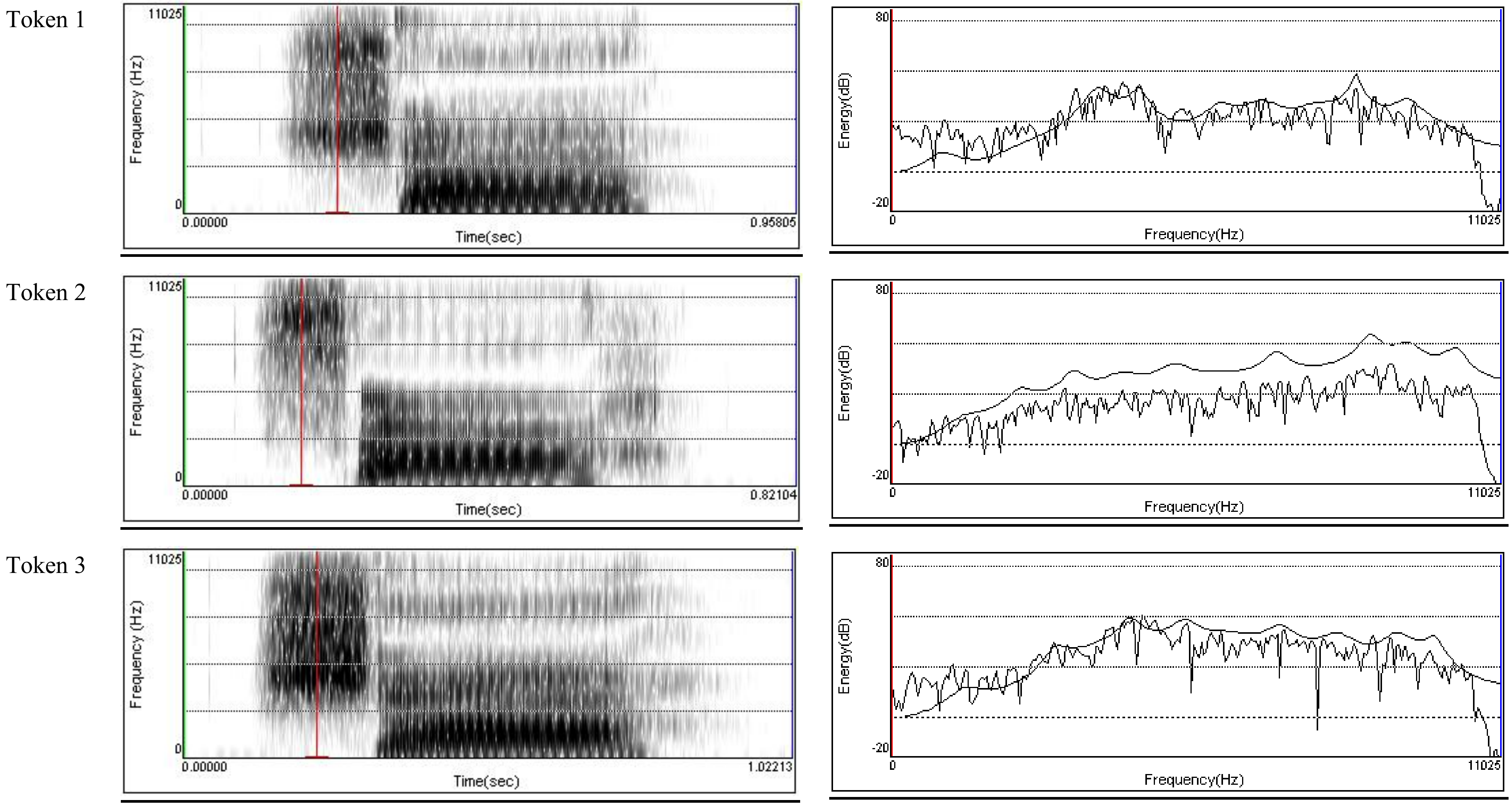


Figure 3.5.6. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [s] in [sa˥] 撒 'cast'

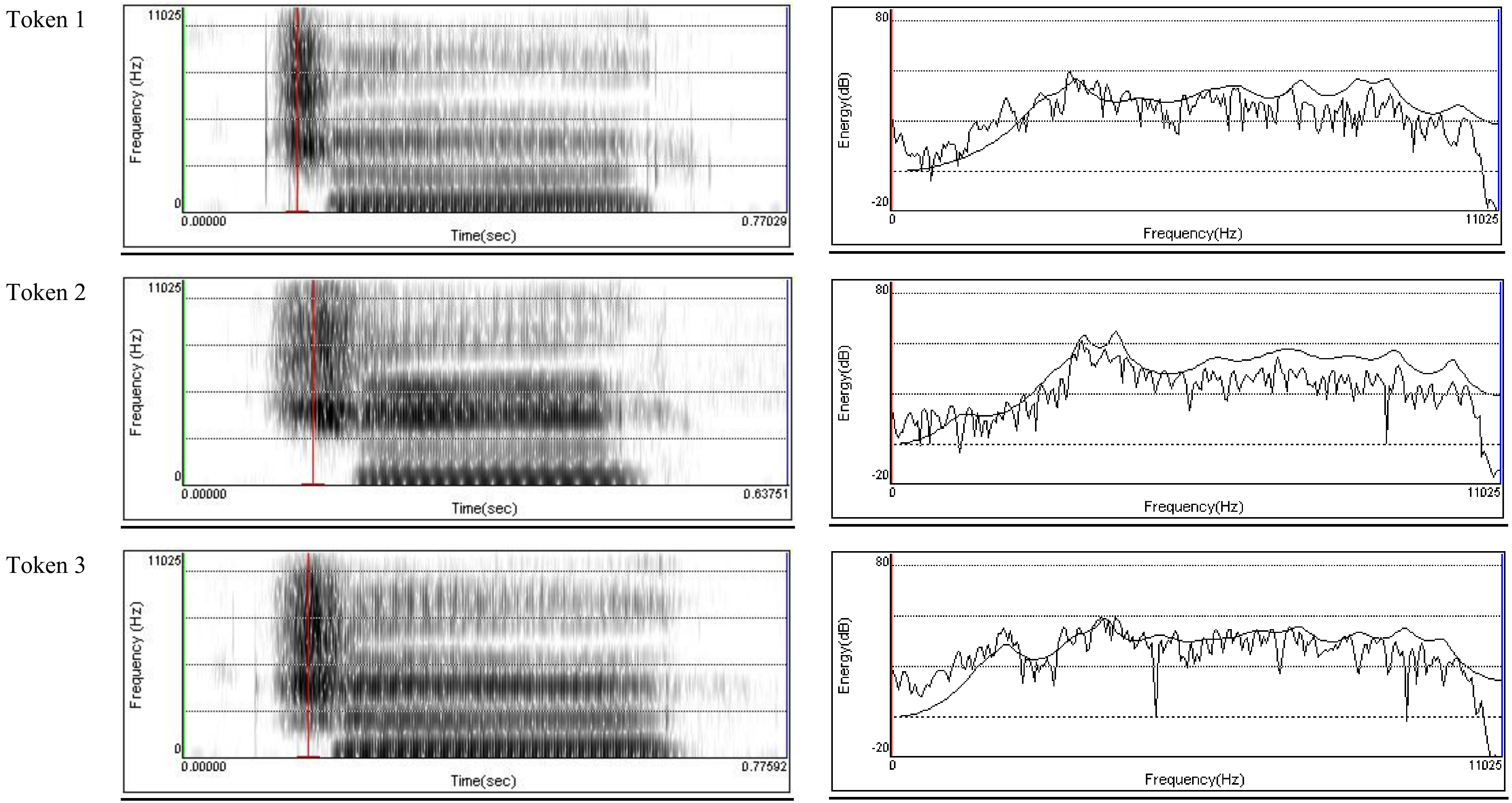
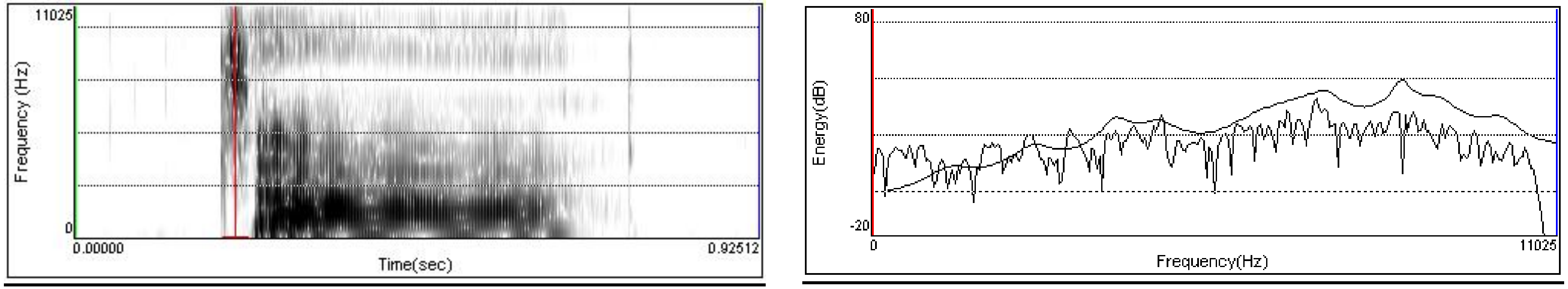
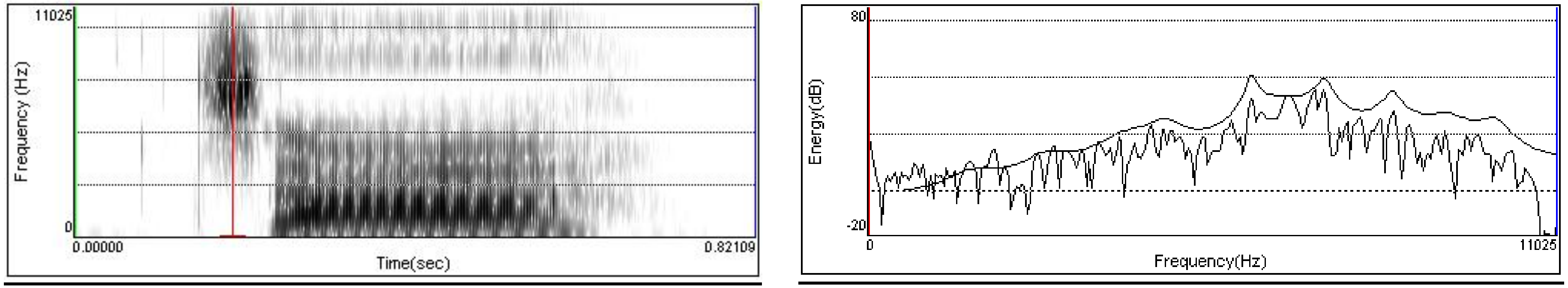


Figure 3.5.7. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [tʂ] in [tʂʅ˥] 知 'know'

Token 1



Token 2



Token 3

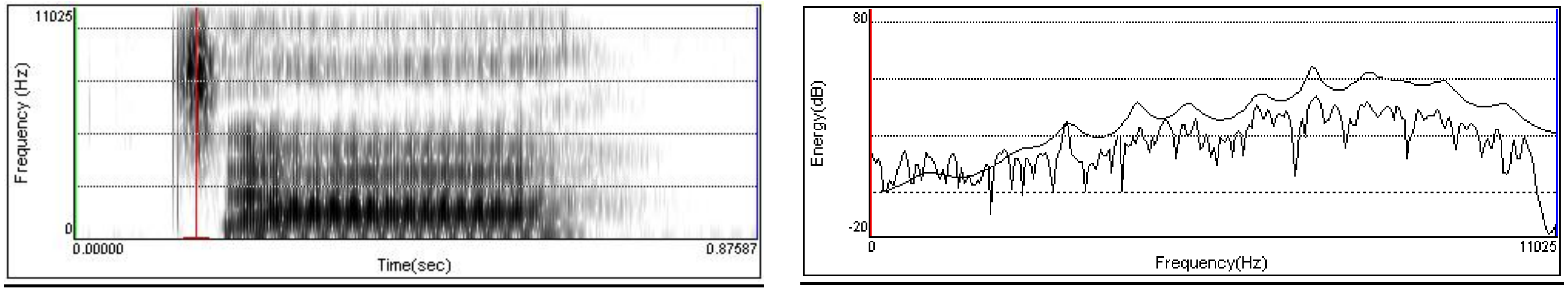


Figure 3.5.8. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [tʂ] in [tʂa˥] 渣 'residue'

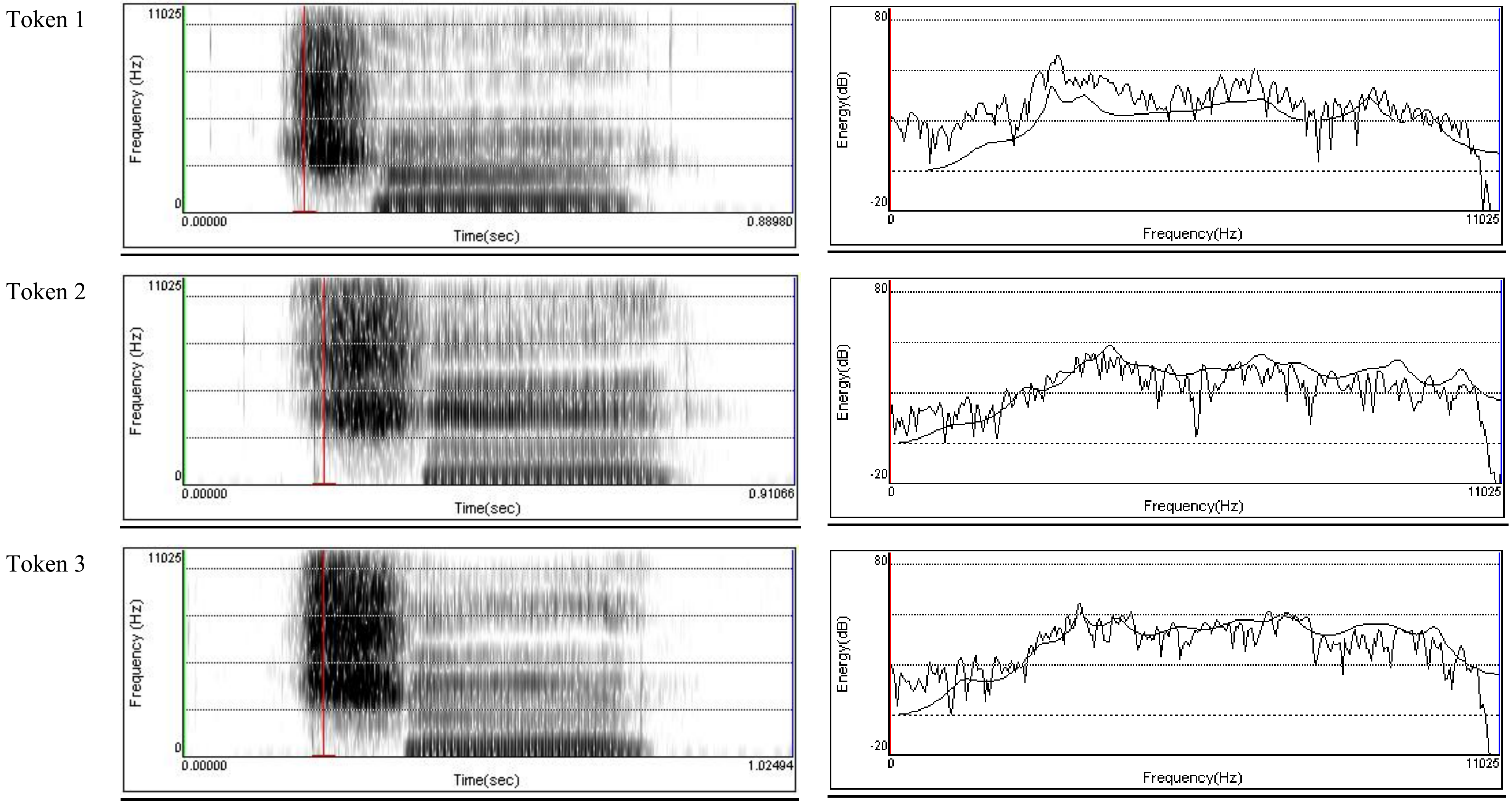


Figure 3.5.9. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [tʂʰ] in [tʂʰʅ˥] 吃 'eat'

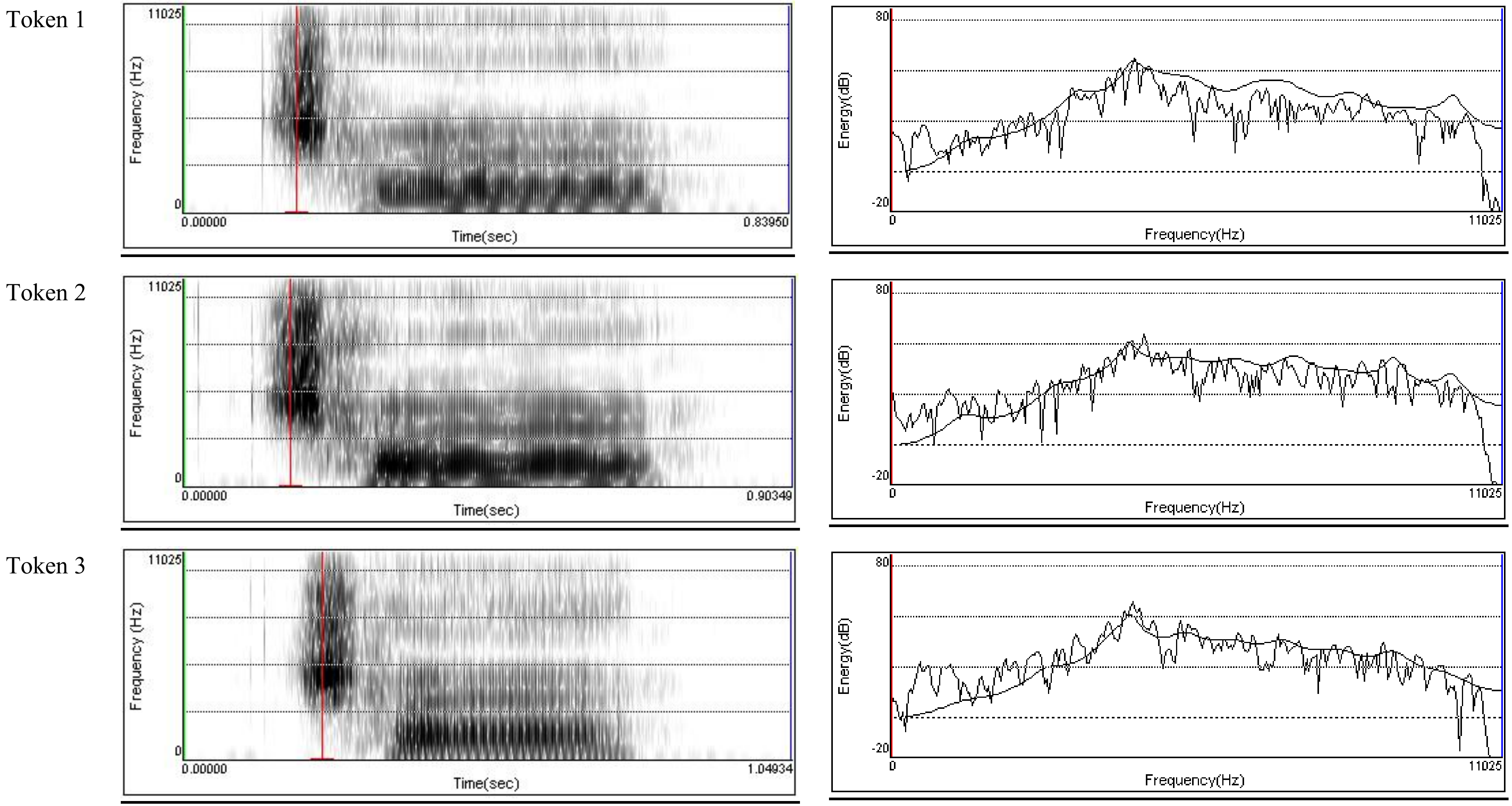


Figure 3.5.10. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [tʂʰ] in [tʂʰa˥] 叉 ‘folk’

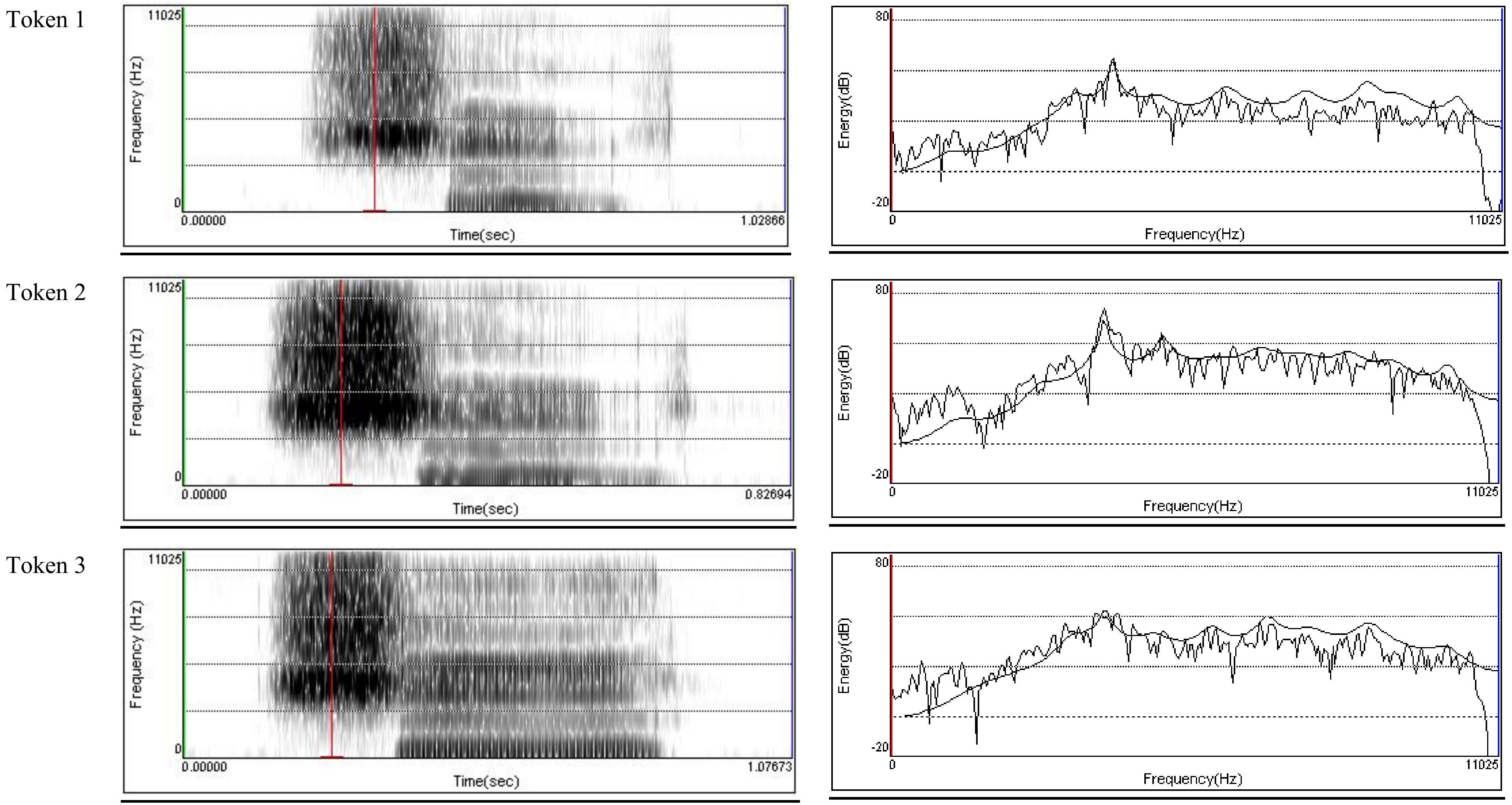


Figure 3.5.11. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [ʂ] in [ʂʅ˥] 失 'lose'

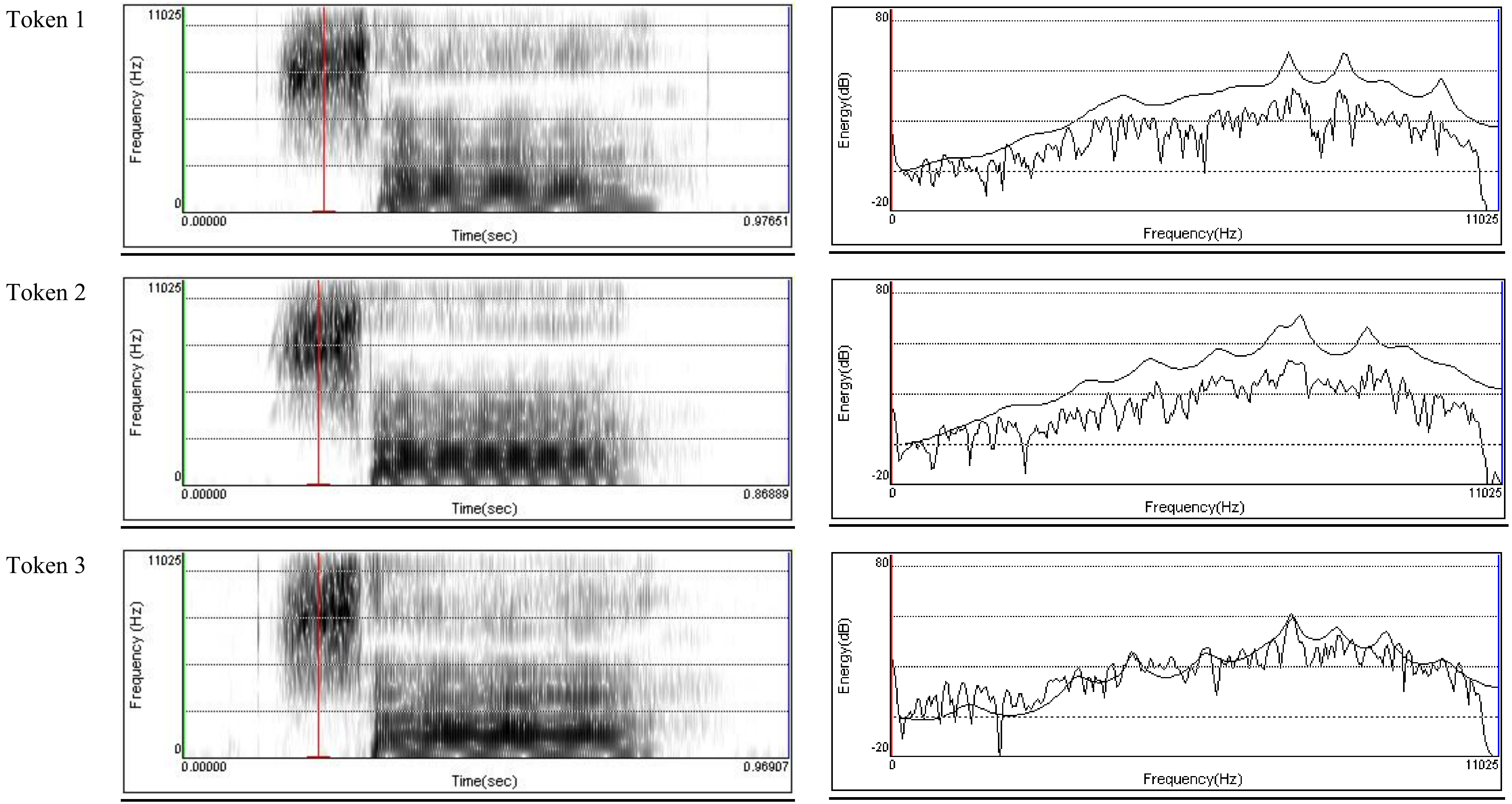


Figure 3.5.12. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [ʂ] in [ʂa˥] 沙 ‘sand’

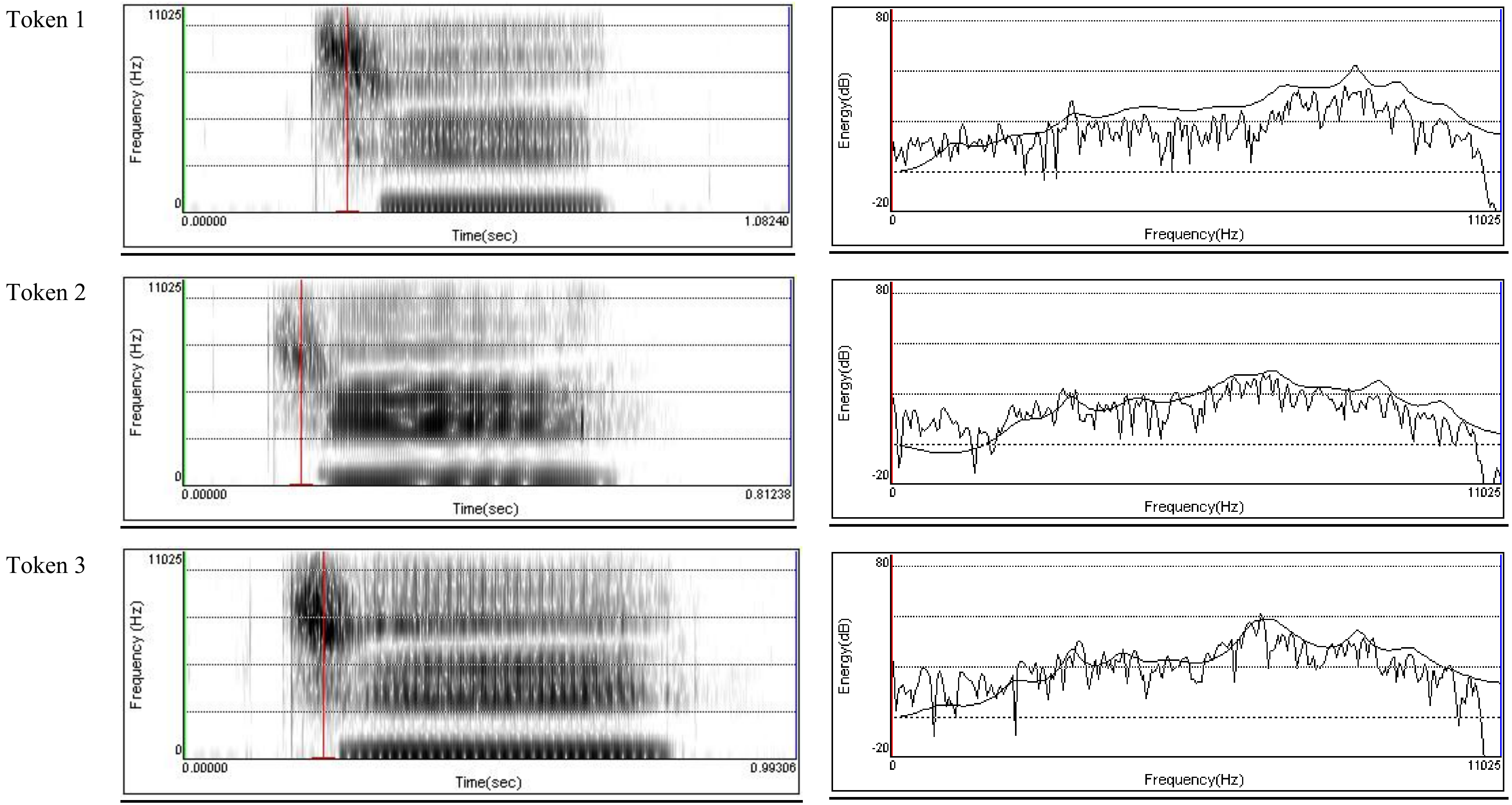


Figure 3.5.13. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [tɕ] in [tɕi˥] 基 'base'

Token 1

Token 2

Token 3

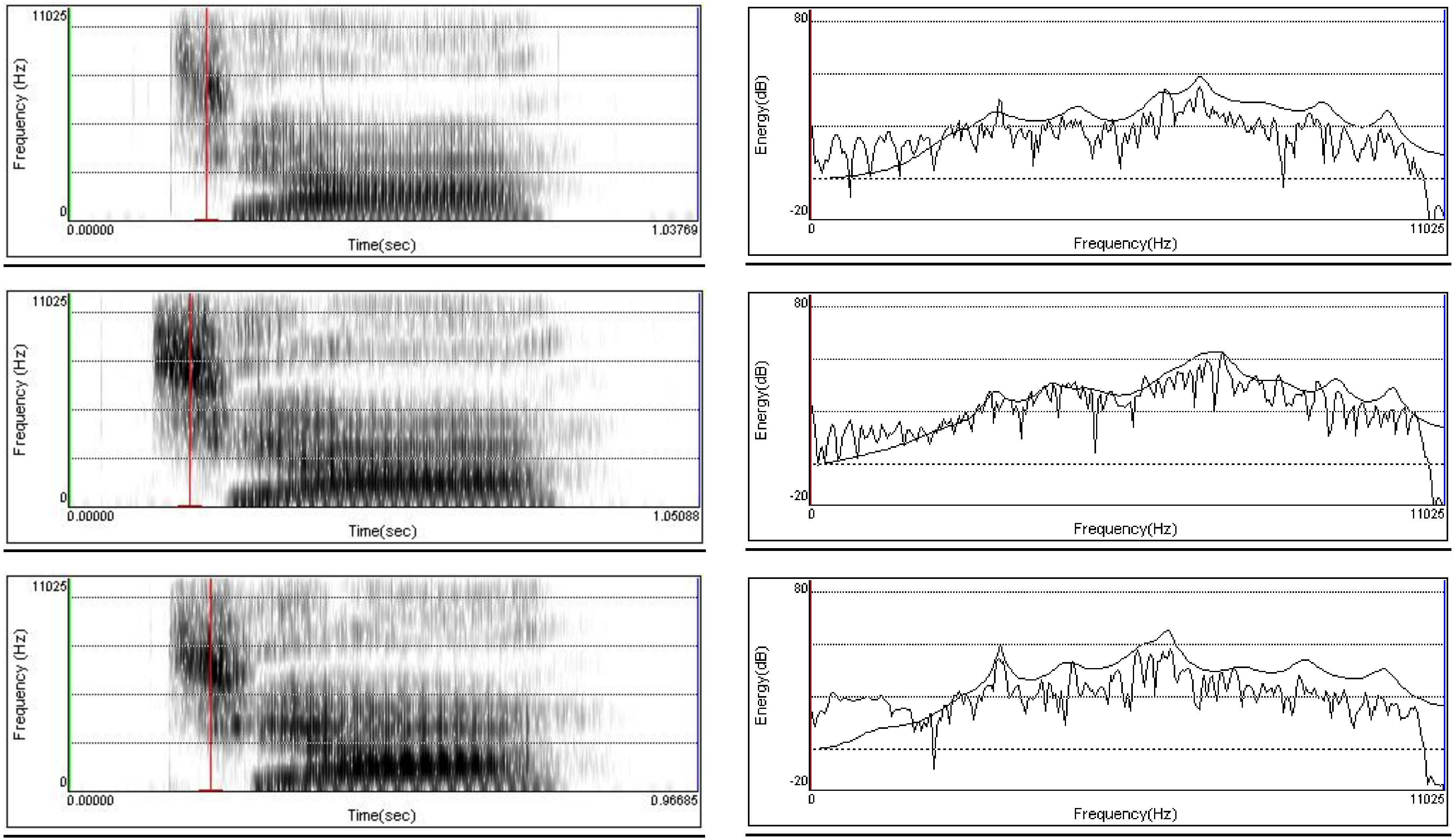


Figure 3.5.14. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the unaspirated affricate [tɕ] in [tɕia˥] 加 ‘add’

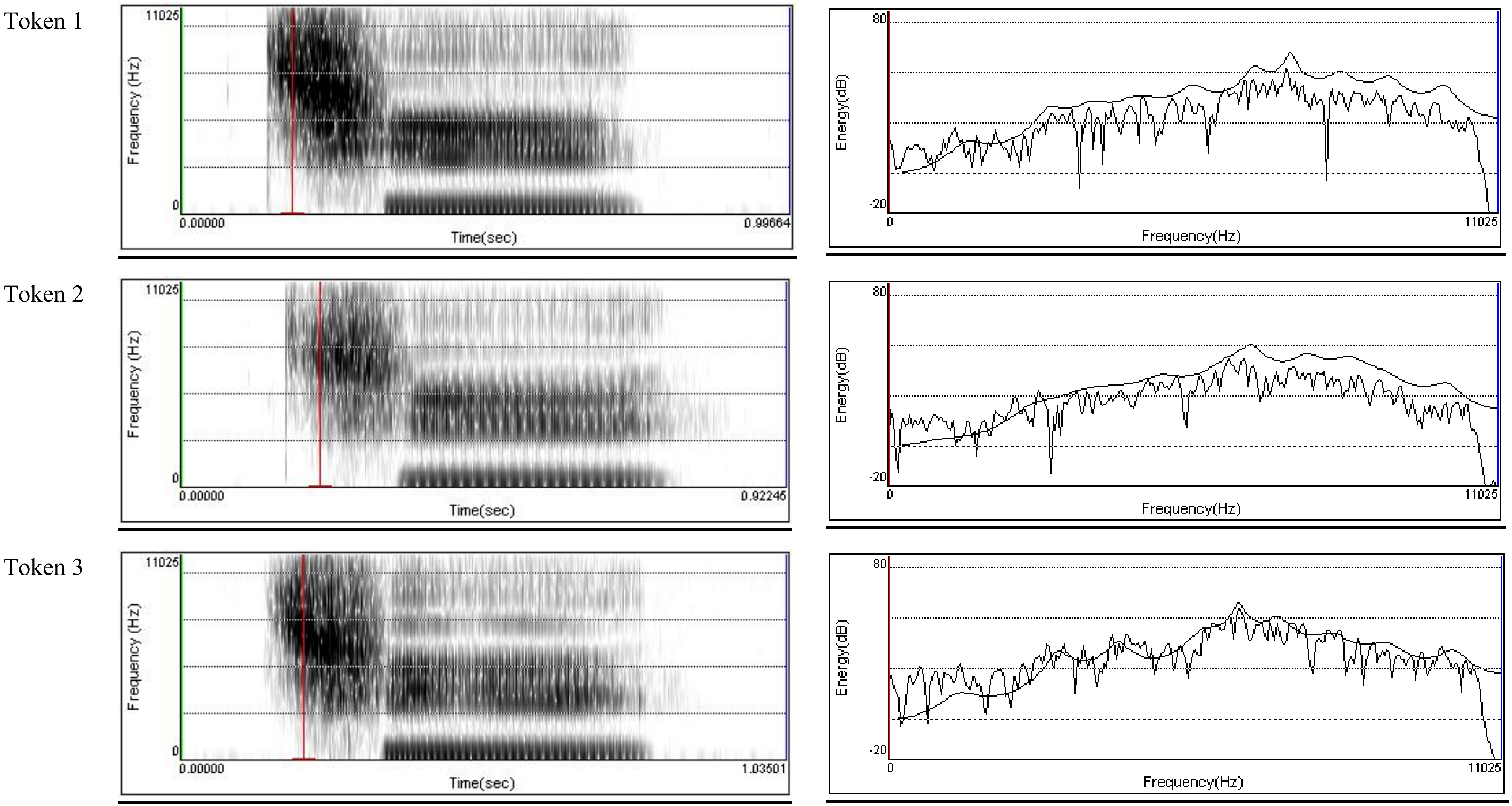


Figure 3.5.15. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [tɕʰ] in [tɕʰi˥] 七 ‘seven’

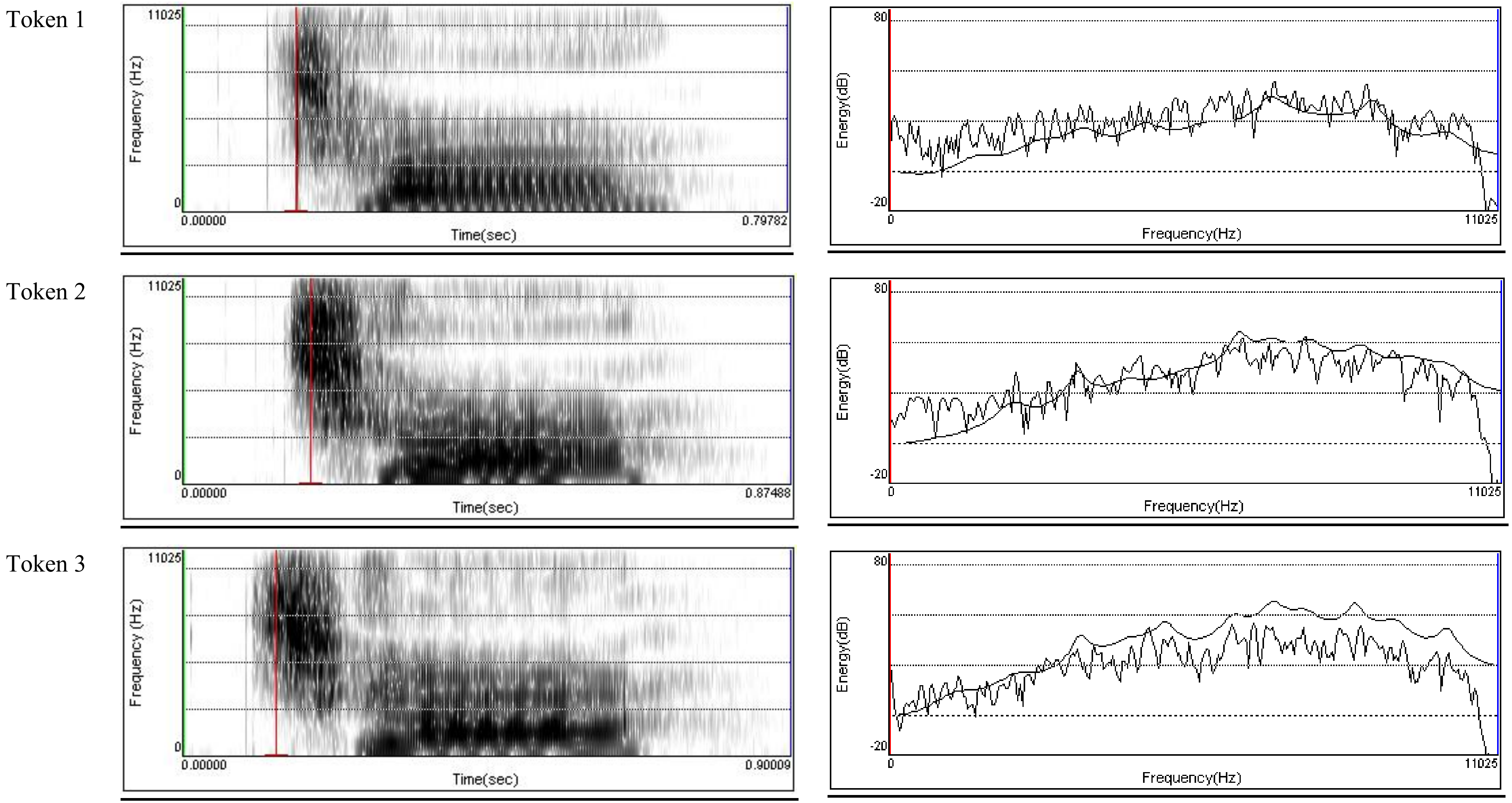


Figure 3.5.16. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the aspirated affricate [tɕ^h] in [tɕ^{h}ia˥] 掐 'nip off'

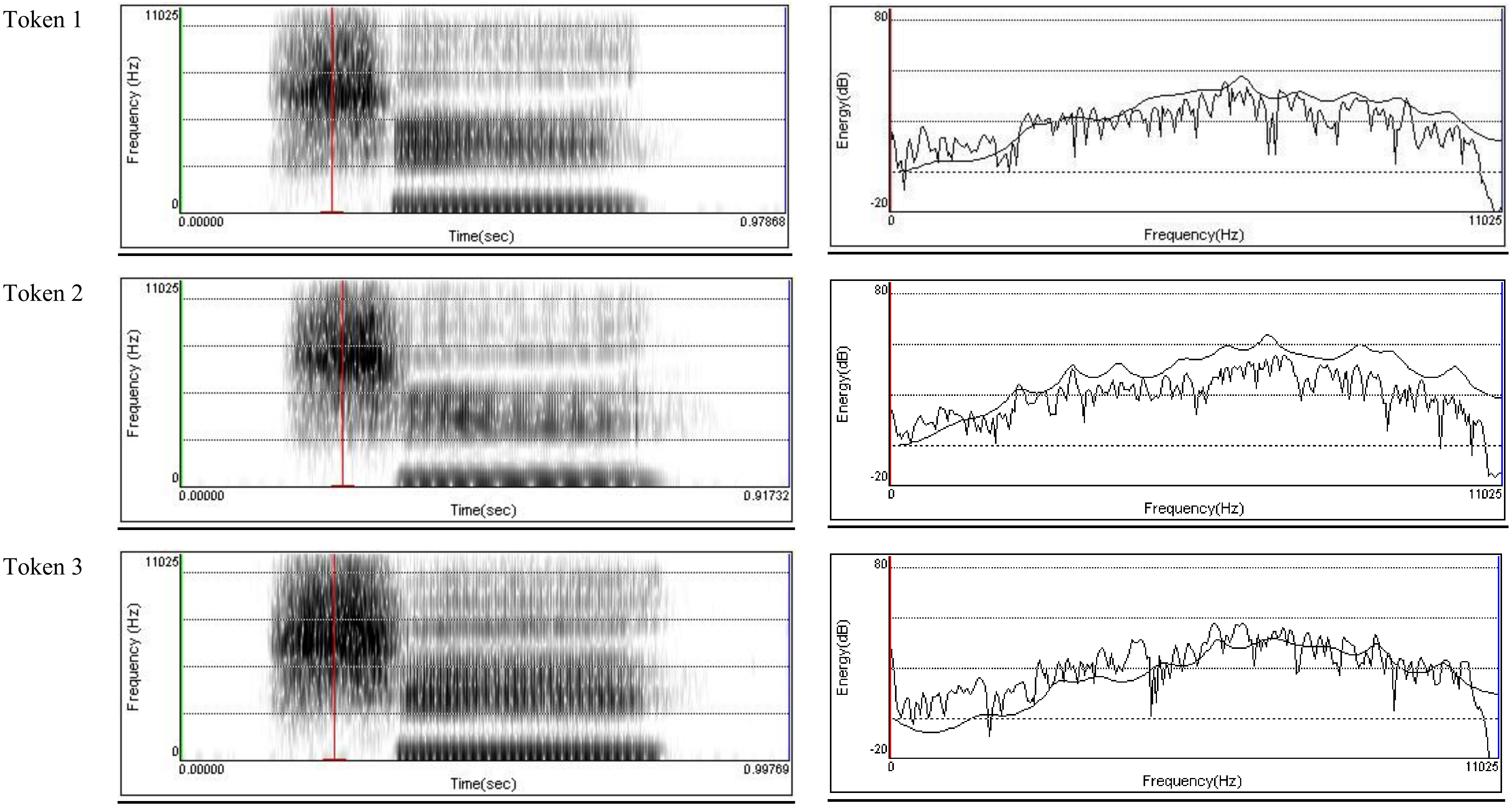


Figure 3.5.17. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [ɕ] in [ɕi˥] 希 ‘hope’

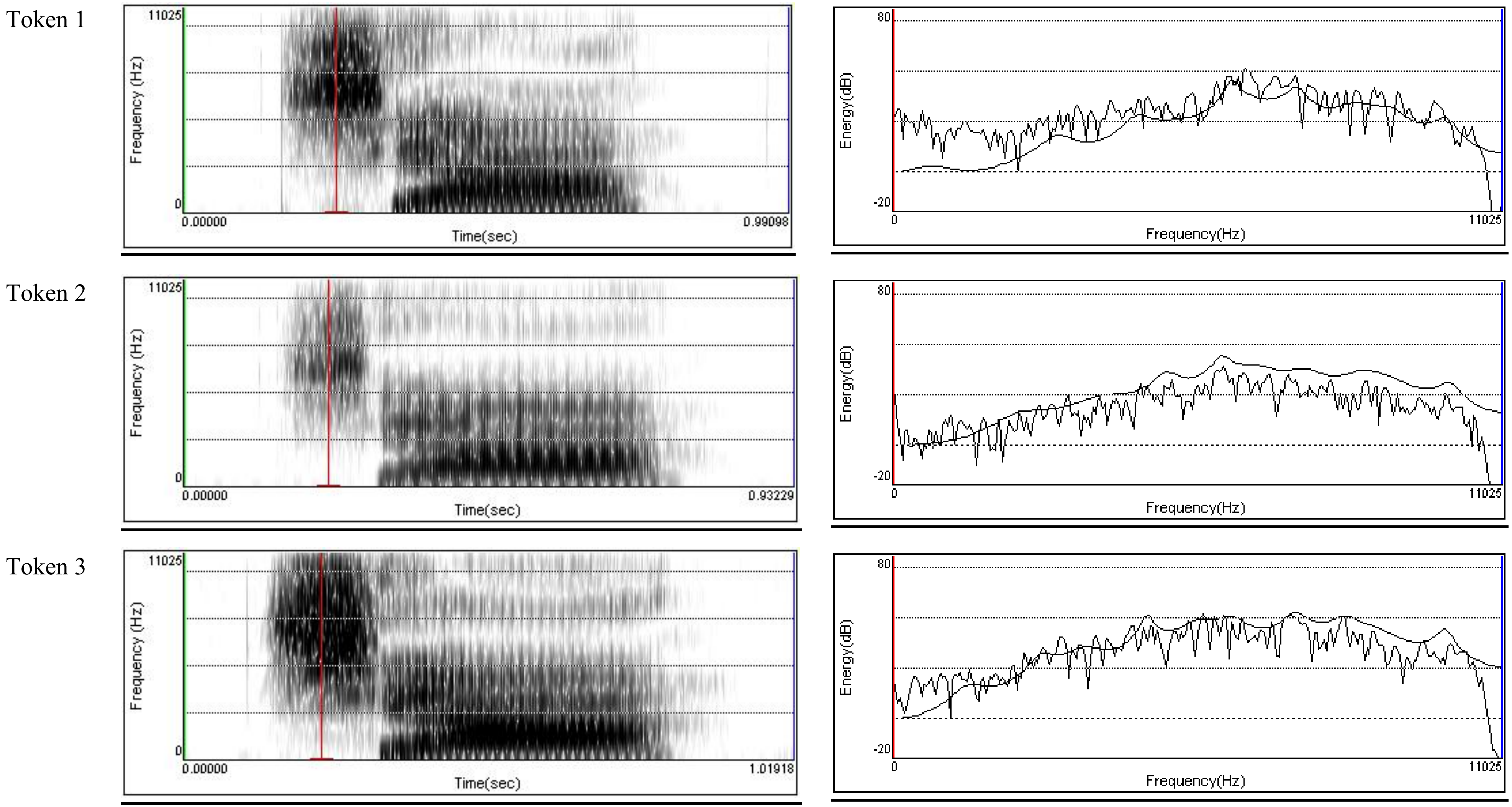


Figure 3.5.18. Wide-band Spectrograms (upper panels) and superimposed LPC & FFT Spectra (lower panels) of the fricative [ɕ] in [ɕia˥] 蝦 ‘prawn’