Tone sandhi, prosodic phrasing, and focus marking in Wenzhou Chinese

Scholz, F.

Author: Scholz, Franziska
Title: Tone sandhi, prosodic phrasing, and focus marking in Wenzhou Chinese
Issue Date: 2012-10-18

Chapter 6

Sentential F_0 scaling in Wenzhou Chinese

6.1 Introduction

6.1.1 Views on pre-planning of sentential F_0 scaling

It has been a long-standing research question in linguistics how much material speakers can (and do) pre-plan before they start uttering a sentence. One way to evaluate the amount of pre-planning from the speaker's perspective is to compare the height of initial F_0 peaks in sentences of different lengths. The basic assumption is that if speakers start higher in longer than in shorter sentences, this means that they take into account the greater number of upcoming syllables and pitch peaks, and adjust their implementation of the intonational melody of the sentence from the very beginning. This idea of an inherent link between sentence length and initial peak height as evidence of pre-planning has been put forward for languages such as Swedish (Bruce 1977; Gårding 1979), Danish (Thorsen 1980), and English (Sorensen & Cooper 1980), and is sometimes referred to as the "global hypothesis" (Prieto et al. 2006).

However, the possibility of an inherent link between sentence length and the scaling of initial F_0 peaks in a sentence has also been challenged on experimental grounds. In particular, Pierrehumbert 1979, 1980 and Liberman & Pierrehumbert 1984 argue that the lowering of successive pitch peaks in English can be modeled more accurately by assuming a constant F_0 decay from one accent to another, and that speakers may either reach their F_0 baseline before the end of the sentence or adjust the slope of the F_0 decline according to sentence length. Such a view of been referred to as the "local hypothesis" (Prieto et al. 2006).

Over the last decades, several studies in different languages have attempted to tease apart the two hypotheses (e.g. Avesani 1987 for Italian, Ladd & Johnson 1987 for English, Kubozono 1993 for Japanese, Prieto et al. 1996 for Mexican Spanish, Rialland 2001 for Dagara, Arvaniti 2003 for Greek, Laniran & Clements 2003 for Yoruba, Connell 2004 for Mambila, and Prieto et al. 2006 for several Romance languages). Most studies, however, have a comparatively small speaker pool (between two and five speakers per language). Results of these studies have also been complicated by between-speaker variation, when for example only a subset of the speakers seemed to use a particular tactic (such as global pre-planning).

These between-speaker variations have sometimes been taken as evidence for the need of a less restrictive theory. For example, Rialland (2001) re-interpreted the mixed results of Kubozono (1993) to indicate that speakers use a mix of local and global strategies for the pre-planning of F_0 scaling in Japanese sentences. Similarly, Prieto and colleagues (2006) interpreted their mixed results for the Portuguese speakers as “soft pre-planning”, and concluded that F_0 preplanning in this language is better viewed as speaker-dependent optional mechanism. What is worth noting is that in most studies on the connection between sentential pre-planning and F_0 scaling, the focus has been on the length of the whole utterance. It is however possible that sentential constituents may also exert an independent effect on the F_0 scaling. The current study sets out to address this issue further by examining how the length of the whole sentence, as well as the length of the subconstituents (such as subject and object), might affect F_0 scaling.

Apart from sentence length, an important factor that has been reported to affect sentential F_0 scaling is the syntactic complexity of the sentence. In English for example, Ladd & Johnson 1987) show that the relative depth of syntactically embedded constituents is correlated with the strength of the prosodic boundary between these constituents, which is reflected in the relative magnitude of downstep between these constituents. Supporting experimental evidence comes from Dutch (van den Berg et al. 1992), Yoruba (Laniran & Clements 2003), and German (Truckenbrodt 2002, 2007; Truckenbrodt & Féry 2003).

What remains unclear is whether sentential F_0 scaling can also be affected by the embedding of subordinate clauses within a sentence in a similar manner. Specifically, sentences which might have comparable length and linear orders of constituents on the surface may nevertheless vary in the complexity of the syntactic embedding structure (for example, a VP vs. a CP being embedded in a sentence, both resulting in the surface word order SVVO). Our second goal is therefore to examine how the complexity of syntactically embedded structures may affect F_0 scaling.

6.1.2 Scaling of F_0 peaks and valleys in Chinese

There have been a few studies related to the issue of sentential F_0 scaling in Chinese, all of which were concerned with Mandarin. For example, Yuan 2004) and Yang & Wang 2002) investigated the declination of the F_0 minima (“baseline”) to illuminate within-sentence lowering. Shih (2000) focused on the relationship between F_0 peaks and sentence length, and reports a ternary split in the realization of the initial F_0 target in short, medium, and long sentences in Mandarin, but concludes that the intermediate differences are too small to be statistically relevant.

However, several other possible factors which might influence sentential F_0 scaling were not accounted for in Shih's study. Apart from the "narrow focus" she induced by prefacing the target sentences with wh-questions, she also varied the names of the subject referents between conditions. This makes it likely that her subjects might have interpreted the subject referents as contrastively focused even in the condition she labeled as "unmarked reading style" (Shih 2000: 247). Therefore, it is difficult to determine from her experiments the exact influence of constituent length on F_0 scaling in Mandarin, without taking the focus confound into account.

More recently, Wang & Xu (2011) took a step further and investigated the effects of several factors on F_0 scaling in Mandarin, while controlling the possible influence of focus more carefully. The sentences in Wang & Xu 2011 are varied in length by adding one or more modifiers before the object noun, but all sentences have the basic structure SVO. Similar to Shih (2000), they report no significant difference between the height of the initial F_0 peak in short, medium, and long sentences. Two problems remain in the design of Wang & Xu 2011. First, all sentences were elicited under one of the different topic and focus conditions. Therefore, it is unclear how sentence length alone might have affected the F_0 scaling in individual words. Second, Wang & Xu (2011) varied the length of the object across conditions, but measured only the F_0 peak of the subject to infer the effect of declination. Therefore, they might have overlooked a more local effect of the different object lengths, e.g. on the first peak of the object. In other words, they confounded different effects of sentence length and constituent length on F_0 scaling.

The current study will investigate another Chinese dialect - Wenzhou Chinese, which exhibits interesting prosodic differences from Mandarin. Wenzhou Chinese is a southern Wu dialect spoken in the city of Wenzhou in Zhejiang Province. It is known for its complex phonology of tone sandhi, the implementation of which is dependent on the intricate interplay between syntactic coherence, prosodic prominence, and length of the constituents (Chen 2000). This chapter will investigate how, on top of tone sandhi, tonal realization (such as F_0 scaling) is conditioned in the language and to what extent, Wenzhou Chinese might differ from Mandarin Chinese in this respect.

6.1.3 Research questions

To test the influence of sentence length and constituent length on F_0 scaling separately, the stimulus sentences in the current experiment varied the number of words in the subject and the object constituent independently. To assess the connection between sentential F_0 scaling and the complexity of syntactic structures, three different types of embedded complements (VP, IP, and CP complements) with comparable sentence and constituent length were recorded. Worth noting is that the focus structure of the test sentences was carefully controlled, so as to exclude the possible confounding effect of focus on F_0 scaling and sentence planning.

The specific research questions are recapitulated in the following:

(i) How does the length of the constituents in a sentence affect F_0 scaling in Wenzhou Chinese? Particularly, how does the length of the constituents determine the scaling of the initial peaks within these constituents, and (how) does it interact with the scaling of the F_0 valleys?

In order to answer the first research question, the length of the stimulus sentences was varied systematically by adding words to the subject and object constituent independently. Looking at the height of the first subject peak will allow us to determine whether only one of the two independent factors (subject length or object length) has an effect on F_0 peak scaling, or whether both factors interact. As for declination of F_0 minima, it will be assessed whether sentence length interacts with F_0 scaling at all, and if so, whether the length of the constituents exerts an independent effect.

(ii) How does the complexity of complex sentences affect the F_0 scaling in Wenzhou Chinese? Particularly, in sentences with similar surface order and length that differ in underlying syntactic complexity, is the difference in complexity reflected in the F_0 scaling?

To address this latter question, two different types of comparison will be made. In a first step, sentences with the surface structure SV(VO) will be investigated, which may contain either an embedded VP complement ($SV(VO)_{VP}$) or an embedded IP complement with unrealized (optional) subject ($SV(proVO)_{IP}$). In a second step, sentences with the surface structure SV(SVO) will be investigated, which may contain either an embedded IP with optional realized subject ($SV((pro/S)VO)_{IP}$), an embedded IP with obligatory subject ($SV(SVO)_{IP}$), or an embedded CP (SV(SVO)$_{CP}$). In both investigations, the goal is to determine whether the syntactic difference is reflected in the F_0 scaling of the tonal contours, e.g. in a steeper prosodic embedding of the more complex structures compared to the less complex structures.

6.2 Methods

6.2.1 Stimuli

The stimulus material was composed in a way that it displayed consistent tonal patterns throughout the sentences. In Wenzhou Chinese, a regular tone change process (tone sandhi) affects both syllables in disyllabic compound words, and changes the tone trajectories to specific tone sandhi contours (Chen 2000). All stimuli sentences were therefore exclusively composed of disyllabic words with the same rise-fall tone sandhi contour.[21] This contour results from a combination of any tone on the first syllable with a dipping tone on the second syllable. The investigation of a tonal contour that covers several combinations of lexical tones allowed for a fairly large database of possible words, and consequently for the composition of semantically acceptable stimulus material.

To investigate F_0 scaling in SVO structures, three lexically different sentences were used for recording, and every sentence was varied between three and seven words in length by adding words to the subject and/or object constituent. This results in 27 target sentences per speaker. The three lexical sentences are given in (1), with parentheses indicating the constituent structure and different lengths of constituents.

[21] This is also the reason why a commonly applied test for downstep vs. declination is not available here, namely to compare the downstepping of HL-tone sequences with that of H-tone sequences. Wenzhou does not have a disyllabic tone sandhi contour consisting of H-tones only, and creating sentences consisting only of monosyllabic words would be highly unnatural.

(1)

a.	(美国	(大学	(同学)))	学习
	mɛi kai	*da ɦu*	*doŋ ɦu*	*ɦu zai*
	American	university	classmate	learn
	(重要	(算术	(公式)))	
	dʑuɔ ʔi	*sɵ jø*	*koŋ sɛi*	
	important	arithmetic	formula	

‘The (American (university (classmate))) learns the (important (arithmetic (formula))).’

b.	(美国	(大学	(同学)))	登录
	mɛi kai	*da ɦu*	*doŋ ɦu*	*taŋ lu*
	American	university	classmate	enroll
	(重要	(大学	(目录))).	
	dʑuɔ ʔi	*da ɦu*	*mu lu*	
	important	university	register	

‘The (American (university (classmate))) enrolls in the (important (university (register))).’

c.	(美国	(大学	(同学)))	翻录
	mɛi kai	*da ɦu*	*doŋ ɦu*	*fa lu*
	American	university	classmate	copy
	(新式	(外国	(音乐))).	
	saŋ sɛi	*va kai*	*jaŋ lu*	
	modern	foreign	music	

‘The (American (university (classmate))) copies the (modern (foreign (music))).’

Additionally, to investigate the effect of syntactic embedding on F_0 scaling, all three sentences mentioned above were also recorded in five different types of complex sentences. For sentences with the surface structure SV(VO) and a sentence length of four words, two different embedding strategies were compared:

- $SV(VO)_{VP}$ with VP complement (embedding verb “to plan”)
- $SV(\text{PRO}VO)_{IP}$ with optional unrealized embedded subject (embedding verb “to promise”)

For sentences with the surface structure SV(SVO) and a sentence length of five words, three different embedding strategies were compared:

- SV((PRO/S)VO)$_{IP}$ with optional realized embedded subject (embedding verb "to promise")
- SV(SVO)$_{IP}$ with obligatory embedded subject (embedding verb "to encourage")[22]
- SV(SVO)$_{CP}$ with embedded clause (embedding verb "to remember")

Examples with the sentence from (1a) can be found in (2).

(2)	a.	同学	计划	学习	公式.	
		doŋ ɦu	*tsẓ va*	*ɦu zai*	*koŋ sɛi*	
		classmate	plan	learn	formula	
		'The classmate plans to learn the formula.'				
	b.	同学	允诺	学习	公式.	
		doŋ ɦu	*joŋ nu*	*ɦu zai*	*koŋ sɛi*	
		classmate	promise	learn	formula	
		'The classmate promises to learn the formula.'				
	c.	同学	允诺	老侄	学习	公式.
		doŋ ɦu	*joŋ nu*	*lɜ dzai*	*ɦu zai*	*koŋ sɛi*
		classmate	promise	nephew	learn	formula
		'The classmate promises the nephew to learn the formula.'				

---

[22] It should be noted that the English translation for the two embedding verbs *promise* and *encourage* does not seem to align with the syntactic analysis given here, which treats the NP intervening between the matrix and embedded verb as "embedded subject". In English, the intervening NP would have to be analyzed as object to the matrix clause in both cases. However, the results of the experiment, as presented in section 6.3 of this chapter, show that the Wenzhou speakers seemed to treat these intervening NPs as part of the embedded clause prosodically. The syntactic analysis of embedding structures in Chinese is notoriously difficult in the absence of overt morphological agreement marking, and different options for syntactic analyses of verbs comparable to those in the present experiment have been put forward for Standard Chinese (Huang 1987; Wei 1997). An in-depth analysis of the syntax of the embedding verbs in Wenzhou lies outside of the scope of this chapter.

d.	同学	鼓励	老侄	学习	公式.
	doŋ ɦu	*ku lɛi*	*lɜ dzai*	*ɦu zai*	*koŋ sɛi*
	classmate	encourage	nephew	learn	formula
	'The classmate encourages the nephew to learn the formula.'				
e.	同学	记得	老侄	学习	公式.
	doŋ ɦu	*tsẓ dɛi*	*lɜ dzai*	*ɦu zai*	*koŋ sɛi*
	classmate	remember	nephew	learn	formula
	'The classmate remembers that the nephew learns the formula.'				

The complex sentences were recorded with single word subject and object constituents, in order to keep the number of stimulus sentences within the experiment to a manageable size.

6.2.2 Speakers

Speakers were all between 20 and 29 years of age (mean age = 23;2) born and raised in the inner-city Lucheng district of Wenzhou. None of them reported to have lived outside of Wenzhou for a significant amount of time within the last 5 years, and they spoke the local dialect with their friends and family on a regular basis. All of them were fluent in Standard Chinese, but had no difficulty reading out aloud Chinese characters in their dialect. None reported any hearing or speech impediments. Nineteen speakers were recorded, 13 of whom were female. Of the 19 speakers, eight recorded the stimuli sentences once, and eleven twice. For those speakers with two recordings, the values were averaged over the recordings before statistical analysis.

6.2.3 Experimental procedure

Speakers were recorded in a sound-proofed recording studio in Wenzhou in individual sessions, and received a small payment for their participation. Each speaker was seated in front of a 13" monitor and given a Sennheiser pc130 headset. The experimenter ensured that the microphone of the headset was placed approximately 3 cm from the corner of the mouth of the subject. Via an external digitizer (UA-G1), the sound was recorded directly on the laptop (Acer Aspire 1810TZ) on which the stimuli were displayed to the subject.

The speakers were first informed about the recording procedure. They were instructed to read out phrases and sentences presented on the screen using Wenzhou dialect in a natural and clear manner. If they were unsure how to pronounce a word or phrase, they could skip to the next item; if they felt they had made a mistake, they could go back and repeat the recording of the previous item. They were told that they could interrupt or abort the recording at any point.

The recording itself was done using a script in the computer program PRAAT (Boersma & Weenink 2001).[23] This script would present the stimulus sentences one by one, and record each stimulus individually after the speaker initiated the recording. Before the recording, all speakers completed a practice series with eight short phrases that were not part of the actual experiment. This was done in order to familiarize the speakers with the self-managed recording procedure. After completing the practice items, the speakers were asked to indicate whether they understood the recording procedure and were ready to start the experiment.

In order to ensure that speakers would not produce the sentences with narrow focus anywhere in the sentence, each sentence was presented together with a precursor question inducing broad focus over the entire sentence (你说什么啊？*ni kuɔ a ni a* – 'What are you saying?'), and the speakers were asked to read out both this question and the target sentence as an answer, as if enacting a dialogue. Furthermore, the target question-answer pairs alternated in the stimulus list with question-answer pairs from another experiment, in a manner that no speaker would see two of the target question-answer pairs from this experiment immediately in a row. This was done to ensure that, in spite of the repetitive nature of the target sentences, the speakers would not interpret the constituents as "given" in the discourse.

6.2.4 Data analysis

After the recording, all utterances were auditorily checked, and if found to contain mistakes and hesitations, excluded from further analysis. The remaining sentences were manually divided into words, using acoustic and visual inspection of the sound wave and spectrogram. A PRAAT script determined the F_0 peak within each disyllabic target word (defined as the F_0 maximum within that word), each F_0 valley between two successive F_0 peaks (defined as the F_0 minimum between two successive peaks) and at the beginning and end of the sentence, and recorded their positions and F_0 values[24]. Before F_0 extraction, the measurements were checked for octave jumps and tracking anomalies, and manually corrected where necessary.

[23] The script used for presenting and recording the stimuli was written by Jos Pacilly, and slightly modified by the author.
[24] The script used for measuring the files was written by Jos Pacilly.

Figure 6.1 shows an example sentence, with the bottom tier indicating word boundaries and the top tier indicating the automatically calculated locations of the F_0 peaks and valleys.

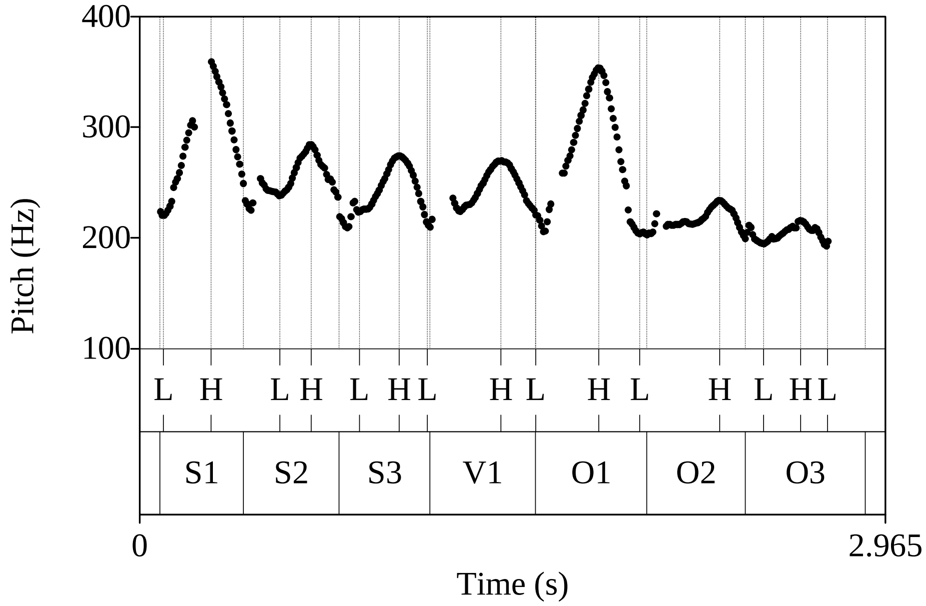

Figure 6.1: Example of a sentence with three subject and three object words, uttered by a female speaker. Dotted lines indicate the location of word boundaries and the automatically determined measurement points for F_0 peaks (H) and valleys (L).

The so determined F_0 values for the peaks (H) and valleys (L) were coded according to the constituent that they occurred on. For example, in Figure 6.1, H within S1 represents the first F_0 peak within the subject constituent, while H within O1 corresponds to the first peak of the object constituent. The so obtained measurements were compared between different combinations of constituent lengths. For plotting after this point, the F_0 values were converted to semitones by applying the formula 12*ln(Hz/x)/ln(2), with x being the pitch floor, which was set to 50 hertz (Hz) for the male speakers, and to 100 Hz for the female speakers. All statistical analyses were performed on the original hertz values.[25]

6.3 Results

6.3.1 F_0 scaling in SVO sentences

The first research question was concerned with the effect of the length of the sentence on the scaling of the F_0 values. First of all, the height of the initial F_0 peak was investigated specifically, because this measure has been most commonly used to investigate the issue of sentential pre-planning. In order to investigate F_0 scaling as a function of the total sentence length, it was pooled over all tested combinations of subject and object lengths, and counted just the words within the sentences. A by-subjects Repeated Measures (RM) ANOVA was conducted with the height of the first subject peak (S1) as dependent variable and WORD number (five levels) as factor.[26] Results showed that the height of the first subject peak was indeed significantly affected by the number of WORDS in the sentence [$F(2.6,18) = 21.17$, $p < 0.001$].

However, in our data, sentence length as a measure is directly related to constituent length, as longer sentences are composed of longer constituents. To investigate the hypothesis that the length of the constituents exerts an independent influence on the scaling of the F_0 peaks, the data was split into the different combinations of subject and object lengths. The first investigation focused on the height of the first object peak and its relation to subject and object length. A by-subjects RM ANOVA was conducted with the height of the first subject peak (S1) as dependent variable and SUBJECT and OBJECT length (three levels each) as factors. Results showed that the height of the first subject peak was significantly affected by the length of the SUBJECT [$F(1.8,14) = 16.79$, $p < 0.001$], but independent of the length of the OBJECT [$F(1.84,14) = 3.0$, $p = 0.071$, ns]. This means that only the number of words in the subject constituent, but not the number of words in the object constituent affected the scaling of the first subject peak, as evident in Figure 6.2.

---

[25] Because the variances in the two speaker subgroups male vs. female are not equal, gender was not included as a factor in the statistical design.

[26] All reported degrees of freedom have been Huyhn-Feldt corrected when the requirement of sphericity was not met.

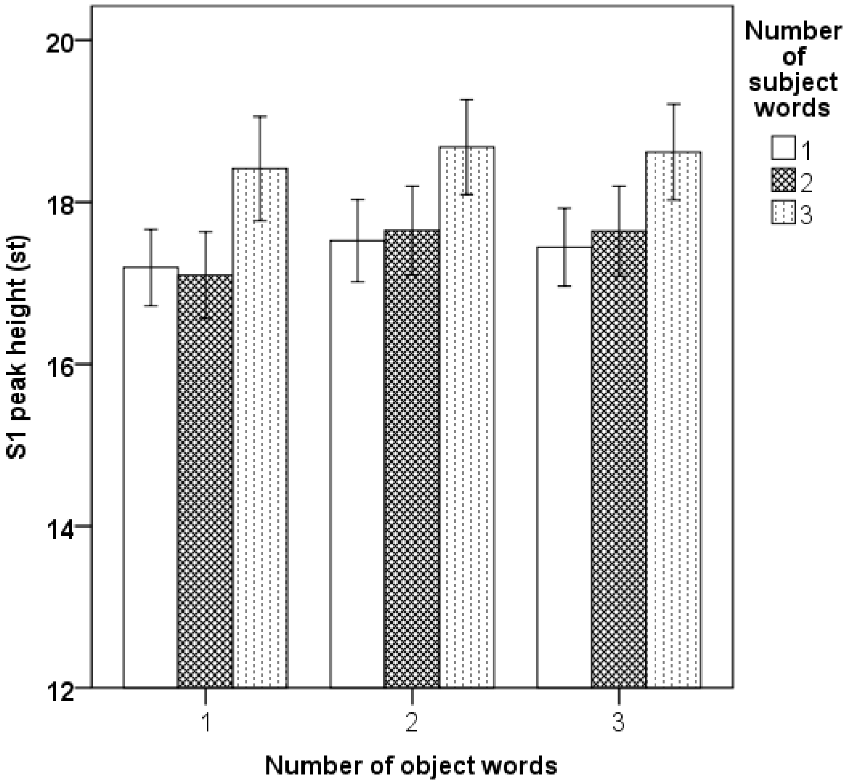

Figure 6.2: F_0 maxima (semitones) on the first subject peak, broken down by constituent length, averaged across speakers. T-bars = ±2 SE.

In order to investigate whether the opposite connection also holds, the height of the first object peak was set in relation to the length of the subject and object constituent respectively. A by-subjects RM ANOVA tested the dependence of the height of first object peak (O1) on the number of words in the SUBJECT and OBJECT constituent (three levels each). Results showed that, across all sentences, the height of the first object peak was dependent on the length of the OBJECT [$F(1.32,14) = 8.11$, $p < 0.01$], but independent of the length of the SUBJECT [$F(1.44,14) = 3.03$, $p = 0.084$, ns]. The results are illustrated in Figure 6.3. Conjointly, these results show that the effect of constituent length on the scaling of the initial peak of that constituent is independent of the length of the other constituent.

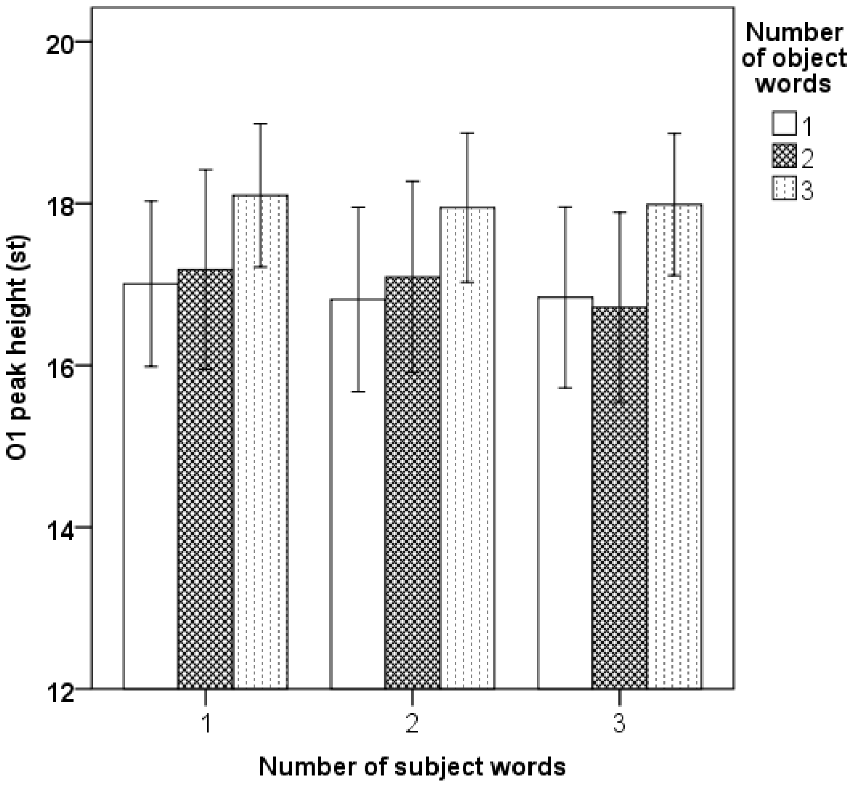

Figure 6.3: F_0 maxima (semitones) on the first object peak, broken down by constituent length, averaged across speakers. T-bars = ±2 SE.

Another interesting observation on the scaling of the first object peaks concerns the question of declination. As can be seen in Figure 6.3, the scaling of the first object peak is independent of the number of words in the subject constituent, even though more words in the subject constituent also mean that the first object peak comes at a later point in the sentence. Apparently, the height to which the first object peak is scaled is independent of its position within the sentence, and of how many words have already preceded it. This is an interesting observation, as it indicates that the scaling of the F_0 peaks in Wenzhou is unaffected by declination.

This finding is further corroborated by investigating the scaling of the F_0 minima across sentences of different lengths. The amount of lowering of the F_0 minima was obtained by subtracting the first from the last F_0 minimum value in each sentence, and then comparing the F_0 minimum ranges across the different sentence lengths (5 levels). A by-subjects RM ANOVA confirms that that there is no consistent increase in the amount of F_0 minima lowering with increased SENTENCE length [$F(3.33,18) = 1.72$, $p = 0.167$, ns]. These results suggest that our findings on the effect of constituent length on F_0 scaling are limited to the scaling of the F_0 peaks, without any declination effect on F_0 maxima or minima.

6.3.2 F_0 scaling in complex sentences

The second main research question concerns the sentential F_0 scaling properties of complex sentences. A comparison of different sentences with similar linear surface order, but different syntactic properties was intended to shed light on this question. First of all, Figure 6.4 illustrates the general tendency that is observed in all complex sentences: The syntactically embedded structures are also embedded in the F_0 scaling, which is lowered on both the F_0 peaks and F_0 valleys after the matrix verb.

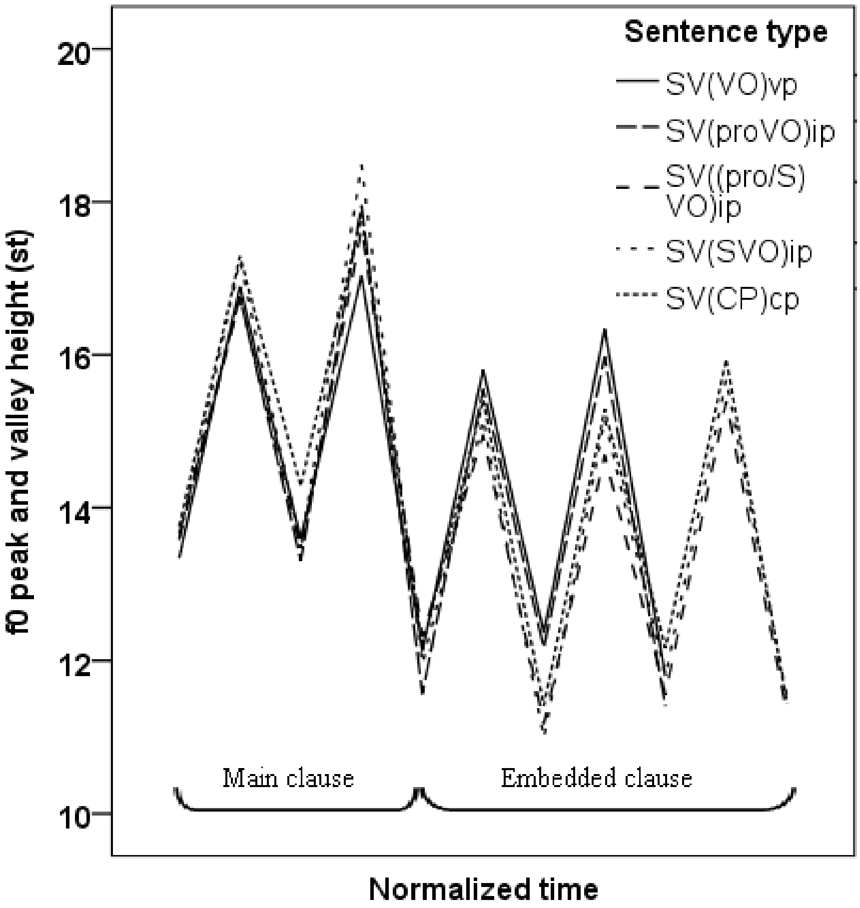

Figure 6.4: Results of F_0 maxima and F_0 minima (semitones) as a function of sequential position of constituents (normalized time), averaged across speakers, for complex sentences with four and five words.

In order to investigate the differences of the F_0 scaling between the sentence types in greater detail, the F_0 difference between the averaged peaks in the matrix clauses and the peaks in the embedded clauses was computed. Secondly, this difference, which represents the average magnitude of drop, was compared across the different sentence types. A by-subjects RM ANOVA showed that the magnitude of F_0 drop in the peaks differed significantly between the five sentence types [$F(3.14,12) = 3.73$, $p < 0.05$]. Post-hoc analysis showed that between the five sentence types, the magnitude of drop for the $SV(VO)_{VP}$-sentences was significantly different from that of all other sentence types ($SV(PROVO)_{IP}$, $SV((PRO/S)VO)_{IP}$, $SV(SVO)_{IP}$, $SV(SVO)_{CP}$), but there was no significant difference between the other four sentence types. The results are graphically represented in Figure 6.5.

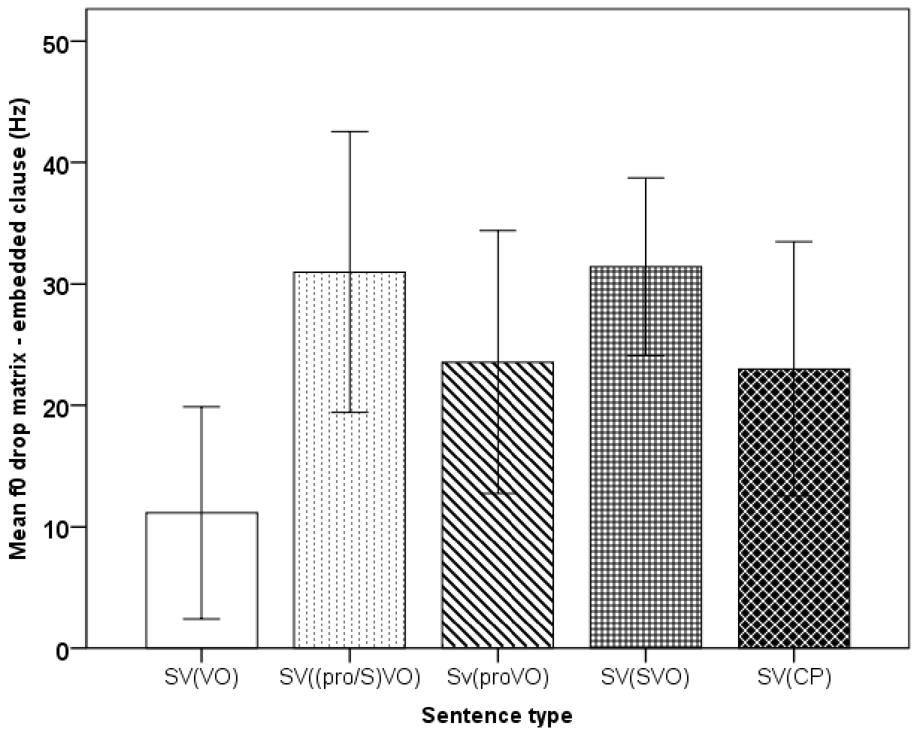

Figure 6.5: F_0 maxima drop (semitones) between matrix and embedded clause, averaged across speakers, for complex sentences with four and five words. T-bars = ±2 SE.

A similar result is obtained when comparing only the scaling of the last peak of the main clause (matrix verb) and the first peak of the embedded clause (embedded verb for $SV(VO)_{VP}$ and $SV(PROVO)_{IP}$, and embedded subject for $SV((PRO/S)VO)_{IP}$, $SV(SVO)_{IP}$, $SV(SVO)_{CP}$), [$F(3.54,12) = 4.04$, $p < 0.01$]. Post-hoc analysis again showed a significant difference between $SV(VO)_{VP}$ and the other sentence types, but no difference among the other four sentence types (SV(PROVO)$_{IP}$, SV((PRO/S)VO)$_{IP}$, SV(SVO)$_{IP}$, SV(SVO)$_{CP}$. This confirms that the F_0 drop is indeed located at the boundary between the matrix and the embedded clause.

6.4 Discussion

This chapter aimed to shed light on two research questions concerning F_0 scaling in Wenzhou Chinese:

(i) Length effect: How does the length of the constituents determine the scaling of the initial peaks within these constituents, and (how) does it interact with the scaling of the F_0 valleys?

(ii) Complexity effect: In sentences with similar surface order and length that differ in underlying syntactic complexity, is the difference in complexity reflected in the F_0 scaling?

In the following paragraphs, the findings of the experiment will be connected to the respective research questions, and discussed with reference to previous research.

6.4.1 F_0 scaling as a function of constituent length

With regard to the first research question, this chapter first examined the connection between sentence length, constituent length, and the scaling of F_0 peaks. Results of our data showed an effect of sentence length on the scaling of the first subject peak. Further investigation, however, shows that the height of the initial F_0 peak is dependent on the length of the subject constituent, regardless of the length of the object constituent. In other words, the observed effect of sentence length is actually an effect of subject length, obscured by conflating the individual contribution of the length of the subject and object constituent. Furthermore, a consistent effect of object length on the scaling of the first object peak could be observed, which is independent of subject length. This finding lends support to the view that the speakers pre-plan the F_0 scaling of sentential constituents one at a time, and that when a new constituent starts, only the length of that constituent is taken into account in the pre-planning of the F_0 height of the initial peak. Preplanning in Wenzhou Chinese therefore is neither strictly “local” nor strictly “global”.

It is worth noting that independent of object F_0 peak scaling, this chapter also found a consistent resetting of the first object peak, i.e. a greater F_0 height on the first object peak than on the preceding verb, regardless of object length.

This resetting is reminiscent of what has been described for e.g. Japanese (Nagahara 1994; Selkirk & Tateishi 1991; Sugahara 2003). What differentiates Wenzhou Chinese from Japanese-type languages is the lack of rephrasing due to the length of the syntactic constituent. In Japanese, it has been observed that speakers opt to vary their prosodic phrasing so as to chunk an utterance into prosodic constituents of similar length and weight, rather than adhere to the syntactic boundaries (Hayashi 2004). Such a weight effect in prosodic phrasing has also been observed in some Romance languages (D'Imperio et al. 2005), and in English (Breen et al. 2011; Watson & Gibson 2004).

This weight effect has also been mentioned in an earlier investigation of Wenzhou Chinese (Chen 2000), but the current data set as presented in this chapter did not lend further supporting evidence. On the contrary, the data reported here shows that for the young Wenzhou speakers, syntactic and semantic coherence is the most important factor for prosodic phrasing. This is reflected in the consistency of the location of the F_0 reset, which occurs on the first object peak (O1) in all conditions. Apparently, the most important information that the speakers want to signal to the hearers is the beginning of a new sentential constituent, and not the approximate length of the upcoming constituent or the sentence as a whole. At the same time, by adjusting the height of the F_0 peaks in accordance to the number of words within the constituent, the speakers still include some information on the size of the upcoming material, while preserving the information of syntactic coherence.

A related issue concerning the scaling of F_0 peaks in Wenzhou Chinese is F_0 declination, i.e. continuous F_0 lowering across sentences. In most previous reports on downstep in tone languages including Mandarin (Wang & Xu 2011), the effects of downstep of F_0 peaks and declination of the F_0 valleys have been described as cumulative (see also Laniran & Clements 2003 for Yoruba). In Wenzhou, it seems that speakers reset their F_0 height to a stable level at the beginning of a new major syntactic constituent, and that this level is affected very little by sentence-internal overall declination of F_0 peak values (cf. Figure 6.3). Similarly, the declination of the baseline values (i.e. low tone targets or F_0 valleys) happens at a stable magnitude across sentences of different lengths. It is therefore independent of the F_0 peak reset discussed earlier, which occurs at a specific point in the syntactic structure, and it is also independent of sentence length.

6.4.2 F_0 scaling as a function of syntactic complexity

In addition to simple SVO sentences, the current chapter also investigated F_0 scaling of embedded clauses within a sentence. Specifically, it addressed the extent to which syntactic complexity is reflected in the intonational scaling of F_0 peaks in complex sentences in Wenzhou Chinese. The data show that the syntactically embedded sentences were also prosodically "embedded" into their matrix clauses, in the sense that both the F_0 peaks and valleys were scaled relatively lower in the embedded clause, compared to the matrix clause. On the surface, this non-local lowering of F_0 targets appears similar to what has been described as "downdrift" or "intonation register" (Yip 1993): A phonological trigger causes all subsequent tones within a prosodic constituent to be realized on an overall lower level. However, in the current study, the trigger for this overall lowering of F_0 range is not phonological (for example a low tone), but structural. The lowering occurs at the juncture of the syntactically embedded clause and its main clause.

Comparing different types of embedded clauses, however, it can be seen that speakers adjust their F_0 scaling to the syntactic complexity of the embedded clause only to a certain extent, even if its surface realization is similar to syntactically less complex structures. Specifically, the matrix subject and verb are realized with significantly higher F_0 peaks when preceding a clausal complement of any type (IP/ CP), than when preceding a VP complement, even when the IP complement has no overtly realized embedded subject. These findings can be interpreted to suggest that, by scaling the matrix structure higher, the speakers anticipate the complexity of the embedded structure and scale accordingly even when certain elements are not overtly realized. At the same time, it is interesting that the scaling of IP complements and CP complements is not very much different from one another.

The findings for complex clauses again underline the important role of syntactic structure for the implementation of F_0 targets in Wenzhou. It appears that it is an important function of F_0 scaling to signal the syntactic relationship between constituents in the sentence, be it clausal embedding (signaled by downshifting the entire F_0 range) or the beginning of a new phrasal constituent (signaled by partial resetting of the F_0 ceiling).

6.4.3 Possibilities for further research

As has been shown in section 6.3.2, the embedded SVO clauses have a similar internal F_0 scaling as the simple SVO clauses, in the sense that the F_0 height of the embedded object is also higher than that of the embedded verb. The beginning of a new phrasal constituent within an embedded clause is therefore also marked with a reset of the F_0 ceiling, within the limits of the overall downshift of the F_0 range due to the syntactic embedding. It can be assumed that the clause-internal F_0 scaling in embedded clauses would be similar in cases with more than one subject and object word, so that the first peak of the subject or object constituent would be scaled higher dependent on the words within that constituent, but still lower than the matrix constituents. With the materials from the experiment reported in this chapter, it would be a straightforward possibility to test this prediction experimentally.

Another interesting possible follow-up research would be to test the influence of other factors that influence the F_0 scaling in sentences, such as focus. From research on Mandarin, it is known that focus has the effect of expanding the F_0 range and consequently raising the F_0 maximum of F_0 peaks, but it is still a matter of debate whether this expansion occurs globally over entire sentences (Shih 1988), or is locally concentrated on the focused constituent (Wang & Xu 2011). Previous research has suggested that prosody and focus may have similar effects on the surface (for example, both may result in a raising of the F_0 peak reference line), but that the two effects are phonetically distinguishable and therefore should not be treated as two sides of the same coin, but as independent factors (Chen 2004; Ishihara 2011; Wang & Xu 2011). For Wenzhou Chinese, it would be interesting to investigate the influence of focus on the F_0 scaling effects reported in this chapter, and to compare the observed similarities and differences to the findings in other languages.

6.5 Conclusion

This chapter investigates the relationship between sentence length, constituent length, syntactic complexity, and F_0 scaling in sentences in Wenzhou Chinese. By varying the length of the constituents in SVO sentences independently of each other, the scaling of the sentence-initial F_0 peak in Wenzhou is shown to be determined by the number of words within the subject constituent, regardless of the length of the object constituent. Likewise, the scaling of the first peak within the object constituent in SVO sentences is solely dependent on the number of words within the object constituent, regardless of the length of the subject.

Within the SVO sentences, the speakers displayed a stable tendency to “reset” the F_0 peak at the beginning of the object constituent. This finding speaks for a syntactically derived prosodic structure, which governs F_0 scaling across sentences in Wenzhou. At the same time, the height of this F_0 reset is resistant to the influence of any general effect of sentential lowering (declination). This suggests that downstep and declination in Wenzhou are not cumulative effects, but that downstep is able to temporarily suspend topline declination.

In complex sentences with embedded VP, IP, or CP complements, the entire F_0 range of the embedded F_0 peaks and valleys was lowered and compressed. At the same time, the relative scaling of the constituents within the embedded clause with respect to each other was similar to that in simple SVO sentences. This leads to the possibility that the lowering of embedded clauses and the reset at phrasal constituent edges are separate processes that occur independently of each other. The reset on the object that occurs within a clausal constituent is constrained by the overall scaling of that constituent in the sentence.

At a different level, a significant prosodic difference between VP complements and clausal complements can be observed, with clausal complements showing a steeper drop of F_0 peaks between the matrix clause and an embedded complex complement, compared to an embedded simple VP complement. In that sense, the speakers are able to mark complex syntactic configurations intonationally, and seem to attempt to give more “room” to the intonational realization of more complex structures. Such a distinction, however, is not present for F_0 scaling of IP vs. CP complements.