COARTICULATION AND GESTURE OVERLAP BETWEEN SYLLABLES

Cao, Jianfen
Institute of Linguistics
Chinese Academy of Social Sciences
5 Jianguomennei Rd.
Beijing 100732
P.R. China
E-mail: Linmc@sun.ihep.ac.cn

Abstract

This paper reports a preliminary investigation on the time course of intersyllabic coarticulation in Standard Chinese. In this investigation, around 3800 phonetically compact C1V1-C2V2 type disyllabic structures are employed to observe the acoustic effect of coarticulation in general, and about 400 disyllabic words are designed as the materials to examine: (1) How the articulators move from one syllable to the next? (2) What is the extent to which the syllables overlapped? And (3) In what sense, the syllables are produced in parallel; and in what sence, they are in sequence? For the convenience of description, we just take the offset transition of V1 and the onset transition of C2 as the rough representations for anticipatory and carryover effect respectively, durational measurements are made correspondingly. To evaluate the possible influence on the behavior of gestural overlap from stress contrast and constituent difference of the syllables, analysis of variance are counducted as well. Based on this study, Some impressions about general nature of coarticulation behined the intersyllabic gesture overlapping in this language are discussed.

1.0 Introduction

Coarticulation exists in all of the fluent speech, it refers to overlap of articulatory gestures associated with adjacent speech segments(Keating, 1988), from which the process of articulatory programing that the segments undergoing in speech production could be observed. Thus, how to explain the phenomenon of coarticulation has become a subject in great demand in the recent 30 years. During this period, many theories and models have contributed, while it still remain as a controvercial issue in this field.The arguments were mainly concentrated on the timing and motor control of speech production. With respect to this issue, different instrumental studies (e.g. Ohman, 1966; Fowler, 1980; Marchal, 1988; Wood, 1991; Zsiga, 1994) related to various segment features have well-attested that overlapping among speech segments is an inescapable nature of speech production. However, most of the studies were dealing with the phenomena occurred in English, Swedish, and Russian etc., but little were involved in the oriental languages, especially for the tone languages such as Chinese. The present study will evaluate whether this nature is the same true in Chinese by examining the time course of articulatory behavior at the boundary between syllables in Standard Chinese. Specifically, the main purpose of this study is to test: (1) how the articulators move from one syllable to the next and whether there exist any information that can mark the syllable boundary? (2) what is the extent to which the neighboring segments may overlapped and does the difference in syllable structure affect the extent of the overlap?

2.0 Materials and Methods

Test materials involved in this study contain two groups of C1V1-C2V2 type disyllabic structures. The first group consists of about 3800 phonetically compact combinations, in which the first syllables contain all the finals and the second ones contain all the initials in Standard Chinese (hereafter SC). These combinations are employed to observe the acoustic effect of coarticulation in general, to see if this effect is occurred commonly in this language. The second group of materials contains about 400 disyllabic words that are devided into two subgroups: the first one consists of 319 normal (NM)stress type words, in which 79 is duplicated ones designed for the convenience of a comparative observation to their spectrograms. In this subgroup, the final (hereafter V1) of the first syllable can be closed by a nasal or a vowel, and the initial (hereafter C2) of the second syllable is selected from various consonants differing either in place or manner of articulation so as to examine possible effect to the manifestation of temporal overlap across syllable boundaries; the second subgroup includes 67 words in neutral(NT) type stress, i.e., the first syllable in the word is normally stressed while the second one is neutralized with weaked stress(Cao, 1992). This group of words is designed to test if different stress status of the syllable have any influence upon temporal characteristics of intersyllabic gesture overlap.

The basis of this study is a careful observation and durational measurements to relevant events through spectrograms, the intersyllabic junctures involved here is represented by V1's offglide transition, C2's onglide transition, and the boundary marker between V1 and C2 . In the case of C2 is voiceless stops or affricates, the C2's on glide is impossible to be observed, so we just take the interval of their closure(gap) as the correlate. The ending point of V1's offglide, which should also be regarded as the starting point of C2 onglide, is defined in terms of the energy end of the second formant and the higher. The reason will be specified below in the section 3.1.1.

3. 0 Observation and Analysis

3.1. *Observation to the acoustic effect*

3.1.1. *Boundary information and measure criterion*

Speech production involves complex motor control. Some previous approaches have claimed that different kinds of segments may go on simultaneously, and thus there are no borders perpendicular to

the time axis in an articulatory or acoustic record to separate one segment from another(Fowler, 1980; Wood, 1991). It may be common in all languages. Thus, with respect to intersyllabic juncture in the SC, a fundamental question facing us is that how to bound the periods of V1 offglide and C2 onglide. To determine the measure points, we have closely examined the spectral effect across syllable boundary in all the test materials, and found that the juncture patterns can be classified in terms of different intersyllabic structures. The examples are shown in Fig. 1, from which we can see that, when V1 is any vowel and C2 is nasal, lateral or voiced fricative[e.g. Fig 1.(1)], the transitions including V1's offglide and C2's onglide is observable and the measure points are easier to be located. Otherwise, especially when V1 is a nasal and C2 is a voiceless stop or affricate, either the V1 offglide or C2 onglide is difficult to be observed . So in this case, where is the ending point for V1 offglide is the crux of the matter.

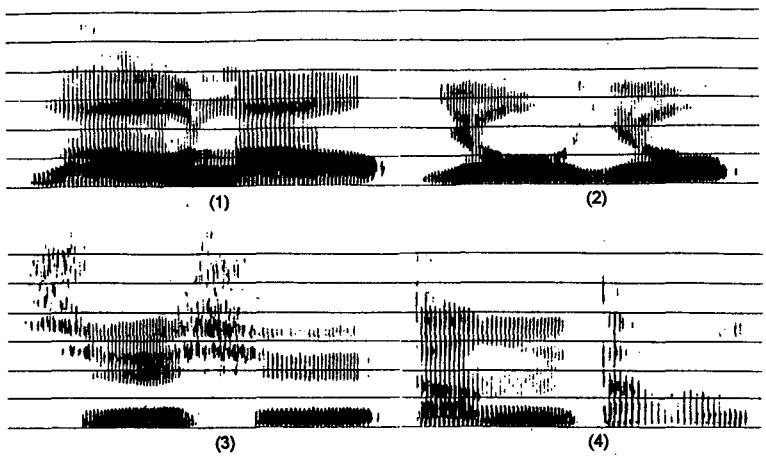

Fig. 1. Spectrographic examples of intersyllabic junctures in Standard Chinese:
(1) /lao-lao/ (2) /iou-iou/ (3) /xu-xu/ (4) /deng-deng/

To overcome this difficulty, let us have a further observation of the case where V1 is a vowel and C2 is nasal or lateral. As what can be seen from Fig. 1.(1), an evident fault block exists between /o/ of the 1st /lao/ and /l/ of the 2nd /lao/, and the formant patterns of transitions are quite different one side from the other of the fault block. Obviously, this fault block is a result of a rapid change in the shape of vocal cavity, It indicates a qualitative variation from /o/ turning to /l/. Consequently, this point should be regarded as the acoustic evidence for the border between the two syllables. A similar evidence is also found from the example shown in Fig.1.(2), where the word is duplicated by two zero-initial syllables, in which the /u/ of the 1st /iou/ is followed by vowel /i/ of the 2nd /iou/, a turning marker between them also can be seen with reference to their formant patterns. Since in both cases, the border occurs right at the F2 end of V1 and the F2 beginning of C2, it is reasonable to take this point as the boundary marker of V1and C2.

3.1.2. *Overlap process*

The observations to acoustic effects described above suggest a physiological process in terms of which the V1 and C2 is coarticulated by following stages of gesture overlapping. Take Fig. 1.(1) as an example, from which we can see that, in the left wing of the fault block between /o/ and /l/, the F2 of /o/ is

rising and the F3 is falling from their target values. It indicates an articulator movement directed towards /l/ has started when the /o/ articulation is on going, so it is a kind of anticipatory coarticulation by overlapping the articulatory tendency of /l/ onto the gestures of /o/. Obviously, this is only a preparing period for /l/, during which the /o/ gesture is still dominant, though it has been interpolated. Therefore, the period before the fault block is actually a process of quantitative variation from the /o/ towards the /l/. On the other hand, in the right wing of the fault block, the formant pattern is essentially belonging to the /l/, but its former portion clearly deviates from its target value. This phenomenon should be caused by a carryover coarticulation, i.e., from the moment of the fault block, the gestures of /l/ have being the dominant, but still to be interpolated by the reminder gesture of /o/, because the gestures of /o/ can not stop at once but withdraw gradually. This kind of mechanism has been revealed by an x-ray approach (Wood. 1991:286) that provides evidence that "the typical utilization of any articulator was that there were brief of movement, and longer periods of inactivity. ...The overall impression is one of the articulators being marshaled momentarily as needed and left alone when not needed, until ..." . According to the proof quoted here, there must be a withdrawal process for the articulators involved in the articulation of /o/. This is also a process of quantitative change taken place during the former portion of /l/.

The general observation to spectrograms of all the test materials found that the acoustic effect of coarticulation in various combinations is identical, it means the mechanism of speech production for various structures must be identical in the same language. Therefore, it is quite reasonable to assume that, the process of temporal overlapping observed from the examples of /lao-lao/ or /iou-iou/ should also be true for the cases other than them.

3.2. *The extent of temporal overlapping*

3.2.1. *General impression*

Durational measurements for relevant events are summarized in the Tables I and II.Taking an overview to the data shown in Table I, a general impression is that, on the one hand, in respect of absolute duration, the mean value for V1 offglide or C2 onglide is quite close to each other in the NM type words, no matter whether C2 is a zero initial or consonant initial. This phenomenon seems to show that, the extent of temporal overlapping between syllables does follow a rough constant, and in the NM type of disyllabic structure, the interval of anticipatory coarticulation is approximated to that of carryover coarticulation. At least, it is true for the situation with respect to the same speaker. On the other hand, however, there is some difference between the figures of subgroup one and two. Besides, a tremendous difference between the minimal and maximal values listed in Table I indicates that, there also exists considerable difference among the individual tokens. These internal differences seem to be caused by multiple factors, which will be specified in 3.2.2.

Table I. The mean duration, sd, durational ratio(to V1 in the 1st sylla. and to C2 in 2nd sylla.) and the minimal - maximal value of the offglide and onglide in NM type(subgroup one) and NT type (subgroup two)words: (1) in the case of C2 is zero initial; (2) in the case of C2 is various consonant initial.

groups	tokens	offglide of V1				onglide of C2			
		mean(ms)	sd	ratio(%)	min-max(ms)	mean	sd	ratio(%)	min-max
subgroup one	(1) 38	69	18	21	45 - 98	66	19	18	38 - 102
	(2) 281	71	15	24	31 - 120	68	17	51	31 - 138

subgroup	(1) 9	60	43	25	46 - 95	32	22	23	13 - 55
two	(2) 58	71	16	25	46 -104	30	15	38	0 - 74

Table II. The mean duration and sd of V1's offglide and C2's onglide in duplicated NM type word: (1) in the case of 1st syllable with a nasal ending; (2) in the case of 1st syllable with a vowel ending; (3) when C2 is unaspirated stops; (4) C2 is aspirated stops; (5) C2 is nasal, lateral or voiced fricative; (6) C2 is voiceless fricative; (7) C2 is unaspirated affricative and (8) C2 is aspirated affricative.

	V1's offglide			C2's onglide		
	tokens	mean(ms)	sd	tokens	mean(ms)	sd
(1)	30	55	12	30	52	12
(2)	49	70	12	49	70	16
(3)	9	63	24	9	75	33
(4)	9	62	24	9	63	24
(5)	12	52	20	12	47	18
(6)	15	64	21	15	64	22
(7)	9	66	14	9	62	11
(8)	9	53	15	9	50	15

3.2.2. *Internal difference and influence factors*

According to the limited materials tested here, the main influence factors seem to be related to the phonetic characteristics of V1 and C2.

First, as what can be seen from the data of group two listed in Table I, the interval for V1 offglide in NT type ones is somewhat equal to their partners in NM type words, while that of C2 onglide is obviously shorter than that in the NM type ones, and the results from analysis of variance indicate that, this difference is highly significant [$F(1,131)=167.46$; $P<0.001$]. It is becouse that in NT type words, the second syllable is neutralized and shortened greatly(Cao, 1992), so the duration of C2 onglide is shortened correspondingly. This phenomenon may agree with such a suggestion that "the movements of gestures are not timed with respect to an external clock but only with respect to the internal stages of some other gesture." (Zsiga, 1991).

Second, if we compare the mean value in the first two lines of Table II, a systematic distinction between the two sets of data can be observed. It shows an influence of the V1 type upon the manifestation of temporal overlapping between V1 and C2. Specifically, in the case where V1 is a nasal ending, both the offglide and onglide is considerably shorter than that for the case where V1 is a vowel, and this kind of influence is statistically significant[for V1's offglide, $F(1,77)=30.72$, $P<0.001$; for C2's onglide, $F(1, 77)=29.53$, $P<0.001$].

Thirdly, the influence from the difference on articulation manner of C2 is also existed. For example, comparing the mean durations listed in lines (3) to (8) of Table II, a slight difference can be observed between the cases where C2 is nasal, lateral or voiced fricative and where C2 is voiceless fricative, as well as between the cases where C2 is aspirated and unaspirated stop or affricates, but no significance has been found here. So far, no clear influence has been found from the difference of articulation place

upon the manifestation of temporal overlapping.

4. Conclusion

According to meticulously observation to the acoustic effects and durational data measured here, we find that the gestures towards C2 always begin during the articulation of V1, and those withdrawing from V1 always end during the articulation of C2, and both of them seem to follow a regular time, though it is considerably influenced by stress contrast and structure difference of the syllables. Based on these facts, we would claim that in speech production, the preparing action for a syllable always appear in a fixed time before the onset of the syllable; likewise, the withdrawing action also complete in a fixed time after the offset of the syllable. Consequently, an overlapped period does exist between adjacent syllables. In this sense, syllables in successive are produced in parallel. On the other hand, however, these anticipatory or carryover effects in articulation does not affect the distinguishing of indivitual syllables in perception. It indicates that such kind of gestural overlapping may only involve the stages of planing and beginning of a syllable , or the latter portion and withdrawing period for a syllable, instead of the period of perception center related to each syllable. In this sense, syllables are produced in sequence. Perhaps, it is more true in tone languages like Standard Chinese, since in this language, each syllable must have its own tone or tone sandhi pattern that closely related to its linguistic function, and it is not allow to be masked in temporal domain. Therefore, neither anticipatory nor carryover coarticulatory effect for a syllable can go beyond the perception center of neiboghring syllable. Of course, the later suggestion is only an assumption, the proper conclusion still can not make until futher study to be conduced.

References

Cao, Jianfen (1992) On neutral-tone syllables in Mandarin Chinese, *Canadian Acoustics*, **20:3**, 49-50.

Fowler, C. (1980) Coarticulation and theories of extrinsic timing, *Journal of Phonetics*, **8**,113-133.

Keating, P. (1988) Coarticulation and timing, In *UCLA Working Papers*, **69**, 1-2.

Marchal, A. (1988) Coproduction: evidence from EPG data, *Speech Communication*, **7**:287-295.

Ohman, S. (1966) Coarticulation in VCV utterances: spectrographic measurements, *J. Acoust. Sco. Am.* **39**, 151-168.

Wood, S. A. J. (1991) X-ray data on the temporal coordination of speech gestures, *Journal of Phonetics*, **19**, 281-292.

Zsiga, E. C. (1994) Acoustic evidence for gestural overlap in consonant sequences, *Journal of Phonetics*, **22**, 121-140.