Duration and the Temporal Structure of Mandarin Discourse

Li-chiung Yang

Spoken Language Research
Department of Foreign Languages and Literature, Tunghai University, Taiwan
e-mail: lyang@sprynet.com or yang_lc@mail.thu.edu.tw

Abstract
Duration is a primary factor in the structural and cognitive organization of discourse. In this study, we investigate pauses and durational patterns in Mandarin Chinese spontaneous conversation, as well as investigate how reliably such elements can serve as boundary-marking predictors across different types of speech corpora, and how language activities are affected by their cognitive correlates. Our results show that pause duration is significantly correlated with specific boundary status and that syllable duration is inversely correlated with distance to phrase end, suggesting that syllable duration is a robust feature in predicting phrase boundary status. Our findings show that duration features are highly informative and provide valuable information on discourse structure and the expressiveness of cognitive state in language communication.

1. Introduction
Language and speech researchers have traditionally been interested in how pauses and syllable duration serve as markers of discourse organization, and to what degree they are reliable indicators of phrase boundaries. With recent increased scientific attention focusing greater interest on the underlying mental and psychological foundations for human behavior, linguists and cognitive psychologists have devoted increased research efforts to study language phenomena as a mirror of internal cognitive processes. From this point of view, pauses are seen as reflections of time needed to activate stored memories or cognitively adjust to new information (Chafe, 1985). In addition, by focusing on the specific cognitive states that pauses are associated with in conversation, including hesitation, uncertainty, and incomprehension, we can hope to achieve a fuller appreciation for their multi-functional nature in communication. A comprehensive characterization of pause use in language also provides important information for speech synthesis systems by enabling more natural sounding speech and better detection of speaker intention.

2. Duration and pauses: motivation, data and approach
2.1. Data, methodology and approach
Data for this research consist of 2 subsections taken from 6 hours of spontaneous Mandarin dialogues in informal conversational settings, as well as 4 children’s stories read by a native adult speaker of Mandarin. The varied data can be expected to shed light on different speech styles and conversational modes and their effects, giving a more complete picture of the use of pauses and duration elements in speech. For this study, we followed the same general principles and methodology as in our work on the durational system of English (Yang, 2004). Data were digitized and segmented to the syllable level, and durational features, including syllable, word, phrase, and pause durations and distance measures, were extracted automatically. From this data, we used a subset of extended selections to focus on duration. Table 1 summarizes the speech time for each of these selected datasets.

Table 1: Data Summary

<table>
<thead>
<tr>
<th>Data Name</th>
<th>Nature of Data</th>
<th>Total Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>MD1</td>
<td>Spontaneous conversation</td>
<td>16 min.</td>
</tr>
<tr>
<td>MD2</td>
<td>Spontaneous conversation</td>
<td>20 min.</td>
</tr>
<tr>
<td>MD3</td>
<td>Read children’s stories</td>
<td>7 min.</td>
</tr>
</tbody>
</table>

For phrase boundary marking for the spontaneous speech data, a 2-level categorization scheme differentiating major and minor phrase was adopted, resulting in 3 types of labels to account for these boundary pauses as well as internal non-boundary pauses. Phrases were segmented as minor or major corresponding to whether the phrase is a subsidiary or tangential part of a larger idea unit. Major phrases correspond roughly to sentences, while minor phrase are clauses and phrases like prepositional phrases, noun phrase, verb phrase, and fragments.

To study variations in syllable duration, syllables were classified by distance to the next occurring pause or phrase boundary, and syllable durations were automatically extracted from the segmented data. Incidence, location, and duration of pauses and syllables were then analyzed with respect to discourse activities such as interruptions, speaker turn, prompting, question and answer sequences, etc., as well as identifiable indicators of cognitive activity, including hesitations, feedback markers of agreement and disagreement, and linguistic markers of certainty, doubt, and uncertainty. Combining the acoustic measures of incidence and duration with the analytic marking of speaker state and interactive activities provided the basis for identifying systematic patterns of cognitive aspects of language form that are mirrored in duration variations.

3. Boundary vs. non-boundary pauses
3.1. Distribution and frequency of pauses
Pauses have been associated with syntactic phrasing as well as with such cognitively associated activities as hesitation, confusion, non-understanding, and reluctance. Syntactically, pauses have been considered as boundary markers of phrases,
with the specific duration associated with different hierarchical levels of phrase structure and with finality of topic or paragraph (Butterworth, 1975; Chafe, 1985; Oliveira, 2000). Phrases in speech have themselves been associated with semantic/syntactic boundaries, or alternately, with idea units or thought segments (Chafe, 1985).

To find out to what extent are pauses used in different types of speech, we calculated the total number of pauses and phrases, and their total time in absolute terms and as percentages. Table 2 shows summary statistics for the 3 sets of data we analyzed.

As shown, we can see that the total time spent on pauses as percentages of total conversation time varied from 14.25% to 35.3%. These results are comparable to the pause ratio range reported for English and other languages, for example, Oliveira (2000) reported values ranging from 0.13% to 0.31% in a set of 17 narratives. In our data the results reflect the different styles of speech as well as different topics. Table 2 indicates a connection between the incidence of pause use as boundary markers and the percent of time pauses constitute of the total time.

One measure of how well pauses function as boundary markers is the percentage of phrases that end with a pause. In our data, MD2 had the smallest percentage of phrases ending in a pause, 42.9%, and also had the smallest proportion of pauses as a percent of total time, 14.26%. MD3 had a remarkable 98.1% of its phrases ending in a pause, and the highest percentage of pause time, 35.3%, of the 3 sets of data. This very high percentage of phrases that end in a pause is very characteristic of read speech. MD3 consists of read speech of Mandarin children’s stories, and the reading is very endearing, but also very structured, with complete ideas encapsulated in phrases and sentences. Planning for topic development is built into the written record rather than having development built into the written record rather than having memory retrieval. These cognitive factors are reflected in the lower pause duration percentage of 14.2% in MD2.

Another measure of how well pauses characterize phrase boundaries is the proportion of pauses that occur at phrase end. If pauses frequently occur within the phrase, they do not unambiguously mark a boundary. Conversely, if pauses are infrequent within a phrase, then the occurrence of a pause is likely to indicate a boundary. Table 2 shows clearly that in the cognitively more demanding MD1, there is a greater proportion of non-boundary pauses (37.6%) than in the other 2 sets, especially by comparison to the read children’s stories.

4. Pause duration and boundary status

4.1. Major, minor, and non-boundary pauses

The strength of pause as a boundary marker is not unambiguous on the basis of occurrence alone, as Table 2 indicates. But pause duration may provide additional information that helps to distinguish boundary pauses and other pauses. Are phrase boundary pauses and phrase internal pauses different enough in duration to characterize boundary and non-boundary marking? Table 3 presents the number of pauses in each category over all of the speakers in the two larger spontaneous conversations and in the read speech data. Table 3 shows that major phrases have the longest pauses on average, with minor phrases considerably shorter, at .33 seconds, and internal pauses having the shortest average duration. Figure 1 presents the average durations of pauses broken out for the 4 speakers in MD1 and MD2, by whether the pause occurs internally, at the end of a minor phrase, or at the end of a major phrase.

![Average Pause Durations by Type](Image)

Table 3: Average Pause Durations by Type

<table>
<thead>
<tr>
<th>TYPE</th>
<th>NUMBER</th>
<th>AVERAGE DUR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major Boundary</td>
<td>665</td>
<td>.493324361</td>
</tr>
<tr>
<td>Minor Boundary</td>
<td>265</td>
<td>.334432075</td>
</tr>
<tr>
<td>Non-boundary</td>
<td>393</td>
<td>.24833842</td>
</tr>
</tbody>
</table>

Interesting, the pattern seen in Figure 1 is consistent with earlier results that we have previously reported for English (Yang, 2004). Phrase-internal pause have the shortest average duration, pauses at minor phrase endings are somewhat longer, and major phrase end pauses have the longest duration.
4.2. Distributions of pause duration

We analyzed the data to find if the consistent patterns of pause type average durations are adequate to characterize pause status. We found that the degree to which this is possible is dependent on the nature of the specific conversation. In Figures 2a-d, we break out the averages of Figure 1 by histograms of pause duration by type for each speaker. It is evident from these histograms that differentiation of pause boundary status is stronger in the conversation of MD1, for both speakers in that conversation. We postulate that this greater distinction among different pause status arises from the more emotionally involved and cognitively more complex nature of that dialogue, so that degrees of finality are more strongly marked.

The difference between internal pause and boundary pause was much greater for the read children’s stories: the average duration for internal pauses for those readings was 0.285 seconds, while the average boundary pause duration was 0.81 seconds. Figure 3 shows that in this set of reading data, there is a clear distinction in duration between boundary and non-boundary pauses. We note the great separation between pause types here, due to the structurally more marked topic presentation, with internal pausing primarily used for introducing new information and for emphasis. There is virtually no overlap in the pause duration histogram, so for this data, pause duration itself characterizes boundary status.

5. Final lengthening and distance to phrase end

5.1. Final lengthening in Mandarin

Final lengthening of syllables at or near the phrase boundary has been one of the key research findings linking duration and phrase boundaries, particularly with read speech, and there has been some debate on whether final lengthening is confined only to the last syllable. In addition, there has also been controversy about where lengthening begins to occur. Our previous research on English syllable duration (Yang, 2004) confirmed that syllables lengthen close to phrase end. Our results in that study also showed that the final syllable before the phrase end is lengthened the most, and that syllable duration is inversely correlated with distance to phrase end.

For the current study, we calculated syllable durations for Mandarin and studied the influence of phrase end status on duration. Our results show that as in English, syllable lengthening also occurs in Mandarin, and that lengthening is in inverse relationship to the distance to the pause or phrase end. Significantly, lengthening is not limited only to the last syllable before a boundary, but begins by at least 5 or 6 syllables from the pause or boundary, confirming that the process in Mandarin is similar to English and other languages. This suggests that there is an underlying cognitive-physiological link to the verbalization of units of thought that is universal across languages.

5.2. Lexical tones and final lengthening

Since the lexical tones affect the duration of syllables in Mandarin, we show the effects of a boundary on syllable duration by lexical tone in Figure 3 for MD1 and MD2.

In both sets of conversational data, there is final lengthen-
internal pauses actually increased much more than for the phrase boundary rises.

We hypothesize that this may be because internal pauses are often the result of exceptional cognitive uncertainty, since they do not function to signal the ongoing flow and structure of phrases, but are concerned solely with the information and affective status of the subject under discussion.

6. Final lengthening and pause interaction

Viewing pauses and syllable lengthening as distinct separate phenomena provides measures of how important each is in expressing cognitive state and in marking boundary status. However, we can also ask how these different phenomena interact with each other. Does the occurrence of a pause affect the lengthening of syllables, or vice-versa? Figure 6 brings out these interactions clearly. From our 2 sets of spontaneous conversation, we recalculated average syllable duration separately for the cases when a phrase ends in a pause and when a phrase does not end in a pause.

In Figure 6 a consistent pattern is found for both minor and major phrases. When a pause ends the phrase, there is less lengthening of syllable duration. When there is no pause, syllable duration is clearly lengthened to a greater degree.

Cognitively speaking, both pausing and lengthening extend the time, so each buys time and also acts to signal the break. This effect is especially clear on the first syllable prior to the phrase end. This data provides convincing evidence for the essential unity of language duration phenomena.

7. Conclusions

In this paper, we have shown that pauses correlated fairly well for phrase and boundary marking, but the strength of boundary-marking through duration varies across corpora. We have found that as in English, the duration of the pause in

Mandarin is also significantly correlated with specific boundary status and that syllable duration is inversely correlated with distance to phrase end. Our results are consistent with previous research and also support the view that cognition and syntax function together to provide coherent development of topic and expression of individual state at both global and local levels. We conclude that pause distribution and duration are essential components of discourse that provide valuable interactive information on both the constraints and expressiveness of cognitive state in language communication. Our findings demonstrate that duration features are a valuable knowledge source and that it is crucial to integrate such knowledge to enhance performance in spoken language systems.

8. References