

Syntactic ambiguity resolution and the prosodic foot: Cross-language differences

CONRAD PERRY

Swinburne University of Technology

MAN-KIT KAN and STEPHEN MATTHEWS

University of Hong Kong

RICHARD KWOK-SHING WONG

Hong Kong Institute of Education

Received: November 30, 2004

Accepted for publication: September 22, 2005

ADDRESS FOR CORRESPONDENCE

Richard Kwok-Shing Wong, Department of Early Childhood Education, Hong Kong Institute of Education, 10 Lo Ping Road, Tai Po, New Territories, Hong Kong SAR, China.

ABSTRACT

In this study we examined syntactic ambiguity resolution in two different Chinese languages, Cantonese and Mandarin, which are relatively similar grammatically but very different phonologically. We did this using four-character sentences that could be read using two, two-syllable sequences (2-2) or a structure where the first syllable could be read by itself. The results showed that when both potential readings were semantically congruent, Mandarin speakers had a strong preference for the 2-2 structure and they preferred that structure much more than Cantonese speakers did. We attribute this to Mandarin having a more dominant bisyllabic prosodic foot than Cantonese. When the 2-2 meaning was semantically incongruent, however, the alternative structure was preferred by both Mandarin and Cantonese speakers. Overall, the results suggest that, in silent reading tasks and semantically neutral conditions, the prosodic foot is generated automatically and can affect syntactic choices when ambiguity arises.

One of the longest lasting and most controversial questions in language research is to what extent people's grammar parsers are innate and what must be learned by such parsers to be functional. At least according to Fodor (1998), if the grammatical mechanism that people use is universal and fully innate, then the same type of syntactic ambiguity in different languages should be resolved the same way, except under special circumstances. Therefore, according to this logic, cross-language differences in syntactic ambiguity resolution provide a challenge to the hypothesis that people use an innate and universal grammar parser, because they suggest that some aspects of grammar may have to be learned and that this learning might be language specific.

Relative clause attachment is one area of grammar where cross-language differences in syntactic ambiguity resolution preferences have been found. In particular, Cuetos and Mitchell (1988; see also Bates, Devescoci, & D'Amico, 1999; Cuetos, Mitchell, & Corley, 1996) found that when comparable relative clauses in Spanish and English could be attached low or high in complex noun phrases, people's preferences differed across languages. Spanish speakers preferred high attachment, whereas English speakers preferred low attachment. This research was extended by a number of authors who examined the same type of ambiguity in different languages. A summary by Fodor (2002) showed that speakers of different languages show different preferences for low and high attachment, with no obvious single predictor, such as language family (e.g., Germanic, Romance, etc.), being able to explain the difference. She suggested that cross-language differences between where breaks in intonational phrases fall, that is, sentence-level prosodic effects, might be responsible for the differences.

Although previous studies examining prosody and ambiguity resolution have typically focused on aspects of prosody to do with sentence-level processing, there are other prosodic structures that might influence people's resolution of syntactic ambiguities. These other types of prosodic structures fall between the level of the syllable and the sentence and form a hierarchy in a similar way as syntactic structures do, although the levels between syllables and sentences in a prosodic hierarchy structure are, of course, different to syntactic ones (see, e.g., Shattuck-Hufnagel & Turk, 1996, for a review of theories of prosodic hierarchies). One thing that is similar between syntactic and prosodic hierarchies, however, is that it is generally assumed that the boundaries of smaller constituents (e.g., prosodic words) do not overlap the boundaries of larger constituents (e.g., intonational phrases). Thus, prosodic words, for instance, do not usually have syllables split across intonational boundaries. Actually what levels there should be in a prosodic hierarchy differs according to different theories, with some levels being well accepted and others only appearing in individual models. In addition, some of the intermediate levels of prosodic hierarchies do not necessarily have obvious perceptual correlates in all languages, and thus determining where they begin and end inside an utterance can be complex and require more than simple perceptual analysis.

One relatively well-accepted part of prosodic hierarchies is the prosodic (stress) foot (e.g., Hayes, 1995; McCarthy & Prince, 1993; Selkirk, 1980). The prosodic foot was initially proposed to explain various patterns of word-level stress in English. Further research has greatly expanded the number of phenomena it has been used to explain. This research has even been done in languages like Mandarin (e.g., Duanmu, 2002), where perceptual correlates of word-level stress are very weak compared to languages like English.

The prosodic foot is a level in the prosodic hierarchy that occurs directly above the syllable. It works as a template that needs to be filled by a certain number of syllables, with only one spot in each foot allowing a syllable to be heavy (which, in languages like English, is typically referred to as stressed). The actual number of syllables used in prosodic feet and their typical distribution differs in different languages. In some languages, prosodic feet need a minimum of two syllables,

but in other languages, monosyllabic feet can be used (see, e.g., Hayes, 1995, for a survey of prosodic feet in different languages).

The structure of prosodic feet and how they are organized allows for predictions to be made about the way words are stressed in different languages (among other things). Thus, for instance, a word like *California* is typically pronounced like CA.li.FOR.ni.A, rather than ca.li.FOR.NI.A or CA.LI.for.ni.a (Hayes, 1995), and such a pattern is not simply random. This is because the syllables must be arranged into prosodic feet, and only one stressed syllable can be found in each foot. Thus, the first four syllables might be organized into two groups (CA.li) and (FOR.ni), where the first syllable in each group is heavy (stressed) and the second syllable is light. The final syllable might also occur in a prosodic foot that takes two syllables (or perhaps a monosyllabic prosodic foot), with the second spot either being filled with a space or the next syllable in the sentence, depending on the surrounding context of the word.

Because the concept of a prosodic foot has been extremely powerful in predicting patterns of data in phonological domains (e.g., McCarthy & Prince, 1993), it suggests that it is a strong source of constraint on people's language processing. In addition, because it is thought to exist at a low level in potential prosodic hierarchies (i.e., before the intonational phrase, see, e.g., Shattuck-Hufnagel & Turk, 1996, for a review), it may be a constituent of prosody that is more likely to affect people's syntactic choices when reading compared to constituents that are further up the prosodic hierarchy. In this case, because it has been hypothesized that prosodic constituents exist in a hierarchy where each level is usually dependent on the levels below, any effects found in higher level constituents entail that lower level constituents must also have been used.

Apart from being a well-accepted constituent of the prosodic hierarchy, one advantage of examining the prosodic foot in reading over some other aspects of prosody is that in some languages, the number of syllables used in each foot is quite predictable (see, e.g., Hayes, 1995). This is important, because if prosodic feet are very predictable, they might act as a stronger source of constraint compared to when they are not predictable. Thus, when there is competition between multiple factors when grammatical ambiguities exist, the extent that prosodic feet dominate other factors may differ across different languages.

PROSODY AND MANDARIN CHINESE

It has been hypothesized that Mandarin Chinese (also known as Putonghua) is a language that has an extremely dominant bisyllabic prosodic foot, or at least some constituent that organizes syllables into groups bigger than one (e.g., Chen's [2000] minimal rhythmic unit). By dominant, we mean that the type of prosodic foot that is typically preferred is bisyllabic (see also, e.g., Duanmu, 2002; Shih, 1986), even when other types of prosodic foot might potentially be used. Feng (2001, 2002) provided a number of arguments as to why this is based on a historical analysis of changes in Chinese words and a linguistic analysis of a number of different

properties of Chinese. In addition, he gives the following example to show just how strong the effect of the bisyllabic prosodic foot is:

“無肺病牛” (p. 113)
wu2 fei4 bing4 niu2 (Mandarin)
mou5 fai3 beng6 ngau4 (Cantonese)¹
none lung sick cow

This string is ambiguous depending on the way the characters are grouped. The grouping can either be in a (無肺)(病牛) [a sick cow without a lung], (無((肺病)牛)) [There is no new species of cow which has lung disease], or (無(肺病)牛) [a cow with no lung disease] pattern, that is, a 2-2, 1-(2-1), or (1-2)-1 pattern. According to Feng (2002), when Mandarin speakers read this, they prefer the 2-2 pattern over the other patterns. He suggests that this is due to the dominance of the bisyllabic prosodic foot in the language. This is a rather strong claim, because in the example, the 2-2 meaning gives the semantically strange “a sick cow without a lung,” whereas the other two readings are more semantically congruent (at least according to native Chinese speakers). If Feng is correct, then it is an extremely important observation, because it suggests that Mandarin speakers prefer to use a bisyllabic prosodic foot over and above other variables such as semantic congruency. This is potentially different to other languages, where it is thought that semantics affects the generation of syntax either through interactivity or revision of semantically incongruous forms (e.g., Boland, 1997; Frazier, 1979). However, to some extent, this may be dependent on the strength of the semantic manipulation in the single example of Feng, as more incongruent forms might cause a stronger semantic bias.

Although Feng (2002) attributes the results of his example sentence to the use of a bisyllabic prosodic foot, there are a number of potential confounding factors. First, three different syntactic/morphological groupings are present. In particular, the 2-2 reading is an A-N phrase, whereas the 1-(2-1) and (1-2)-1 structure are A-N-N phrases. If people prefer the simplest structure when ambiguity exists, as suggested by the garden path theory of Frazier (1979), the 2-2 structure might be preferred for that reason. However, such a preference would also suggest that no revision is done based on a preference for the most semantically congruent form, which is not predicted by the garden path theory.

Second, the meaning people choose may be influenced by word-form frequency. In this case, because the word frequencies in the 2-2 form are different to the 1-2-1 forms, people may have chosen the 2-2 structure because of some function of word frequency, rather than word prosody. The idea here is that there must be some form of nonlinear relationship (currently not well known) between word frequency and word length, because single syllable words can be of higher frequency than bisyllabic words, but can also be embedded as morphemes within bisyllabic words. Thus, if selection was based purely on absolute frequency comparisons and if ambiguity existed, people would almost never choose bisyllabic compared to monosyllabic words when bisyllabic words have high-frequency monosyllabic morphemes embedded in them, which is unlikely to be true (this general problem is known as the masking problem; see Cohen & Grossberg, 1986).

Third, it is certainly possible to think of ad hoc sentences of a similar form where people do not seem to prefer the bisyllabic breakdown, for reasons that are difficult to immediately ascertain. Thus, the generality of the single sentence presented needs to be extended.

Because there are many potential factors that might influence the way people break down four-character strings, actually determining whether people use a 2-2 breakdown due to a bisyllabic prosodic foot or another reason is somewhat difficult, because it would require a number of factors to be manipulated and examined separately and in conjunction with each other. Because of the lack of potential four-character sentences that have the property where they have ambiguous meanings that allow grammatical and semantic variables to be manipulated, a thorough within-language investigation may be difficult using this type of method, which is unfortunate, because it is a very effective way for investigating the extent that people prefer 2-2 or 1-2-1 patterns.

We should note that it is extremely difficult to cleanly separate syntactic groupings from morphological groupings in Chinese, particularly for new groupings of morphemes, and people argue about the distinction between words and morphemes a great deal (e.g., Bates, Chen, Li, Opie, & Tzeng, 1993; Chao, 1968; Packard, 2000). We will therefore refer to effects where it is difficult to differentiate between syntactic and morphological groupings as hierarchical groupings (which are basically all of the results reported in this paper). By this we mean a prosodic effect on the grouping choice of syllables, whether this choice is due to syntactic choices among words, complex morphological grouping choices, or both.

PROSODY AND CANTONESE

One way around the problem of manipulating individual variables within Mandarin Chinese to examine prosodic effects is to examine other Chinese languages² that have a difference in the dominance in the use of bisyllabic prosodic feet, but are otherwise relatively similar (in the same way as Spanish and Italian or English and German might be considered relatively similar). One such language is Cantonese. According to Bauer and Benedict (1997), compared to Mandarin, Cantonese is similar grammatically but quite different phonologically. In terms of grammar, this is particularly true of short simple phrases that do not involve particles (see Mathews & Yip, 1994, for a description of Cantonese grammar). This similarity also extends to other aspects of the language, such as word form and word frequency. Thus, many words that have the same meaning are likely to have relatively similar frequencies. For instance, it is likely that many common words with direct one to one translations, such as the words for “dog,” are used at a relatively similar overall frequency in Cantonese and Mandarin, in much the same way as they would be for other similar languages. (Thus, e.g., we would expect the frequency of the word for dog to be similar in both Spanish and Italian.) This similarity means that sentences can be constructed that are very similar with respect to word frequency and syntax, but are very different phonologically. Therefore, the effect of prosody can be examined using similar sentences in Mandarin and Cantonese, with potentially confounding variables implicitly controlled. There are, of course, words where the frequency is likely to be different (just as there are

between Spanish and Italian). However, for the experiments reported below, we tried to avoid them.

Of course, for prosodic differences at the level of the prosodic foot to emerge between Cantonese and Mandarin, there would need to be differences in the dominance of the typical prosodic foot used. Although it seems clear that Mandarin speakers generally prefer to use a bisyllabic prosodic foot structure (e.g., Duanmu, 2002; Feng, 2002), the preference in Cantonese is not so clear. For instance, based on her perceptual intuitions, Flynn (2004) claims that Cantonese uses rhythmic foot structures where one, two, and three syllable feet are commonly found. Alternatively, Wong, Chan, and Beckman (2004) offer a different suggestion, claiming that there is a preponderance of monosyllabic feet in Cantonese. Finally, Yip (1993) has suggested that a bisyllabic prosodic foot structure may be commonly used in some circumstances in Cantonese. Unlike Flynn or Wong et al., Yip based her suggestion on observations not open to her own perceptual bias and on factors that do not necessarily have a rhythmic perceptual correlate.

Despite the rather reasonable suggestions of Yip (1993) that Cantonese may use bisyllabic prosodic feet in some circumstances, the extent that bisyllabic feet are used and their grammatical effect may be less than Mandarin, given the extremely constraining nature that bisyllabic prosodic feet appear to have in Mandarin (e.g., Feng, 2002; Duanmu, 2004). Thus, when different forms of information (e.g., semantic, syntactic, prosodic) are in conflict, the preference for bisyllabic feet might be less than in Mandarin. We also note informally that when we give Cantonese speakers the same sentence as Feng (2002; it uses four characters that have one to one translations into Cantonese), they prefer the semantically plausible 1-(2-1) meaning. Therefore, if we are correct in the belief that the bisyllabic prosodic foot is less dominant in Cantonese than Mandarin, then the effect may be observable in terms of the type of meaning that people extract from ambiguous sentences.

Of course, the only way to determine whether Cantonese and Mandarin differ with respect to how their prosodic feet are used is to test whether this is so. Two ways this could be done would be to suggest grammatical tests that distinguish between the two languages, such as those proposed by Duanmu (2002) for Mandarin, or to experimentally test the results. The first of these is simple to do, because of the tests Duanmu uses to examine prosodic foot structure in Mandarin, many are equally as applicable in Cantonese. These include restrictions on word length in verb-object phrases, restrictions on word length in modifier-noun compounds, and preferred synonyms in “de” usage (*ge* in Cantonese). At least according to our informal observations, the results of these tests in Cantonese appear to display a relatively similar pattern, suggesting bisyllabic prosodic feet are preferred in both languages, and hence, they do not distinguish between the two languages well (see Duanmu, 2002, for further information with regard to these tests, and Feng, 2002, for an alternative explanation of some of them; Cantonese examples we have examined that produce similar results to Mandarin include verb-object: [zung3zik6] [faa1deo2] [to plant] [flowers]; modifier-noun: [lou5fyu2] [maa5ai5] [tiger] [ant]; and *ge* usage restrictions: [waai6jan4] ge3 [hei1pin3] [bad-person]’s [cheating]).

Although it appears that the tests proposed by Duanmu (2002) do not distinguish whether the prosodic foot structure differs qualitatively between Cantonese and

Mandarin, there may be less obvious quantitative differences. What we mean by this is that if there are multiple constraints in how people process sentences (e.g., syntactic, semantic, phonological), then the weight of the prosodic foot compared to other constraints may be different. For instance, if there was a tradeoff between a typical prosodic foot structure and a semantically congruent structure, the typical prosodic foot structure might be chosen in Mandarin but not in Cantonese.

CANTONESE AND MANDARIN

Because the strength of the bisyllabic prosodic foot in Cantonese might be different to Mandarin, the difference may allow us to examine whether prosodic differences in languages can cause ambiguities to be resolved differently in silent reading tasks. More specifically, the goal of the following experiment was to examine, using ambiguous four-character strings, the extent that Cantonese first language speakers choose a 2-2 sentence breakdown over a 1-(2-1) sentence breakdown compared to Mandarin first language speakers. (We did not test 2-2 vs. [1-2]-1 sentences because of a lack of potential stimuli.) It is also possible to examine the same idea independently using stimuli with a 2-2 versus a 1-3 grouping; thus, we also used such a group. The idea with both groups is that if people prefer the simplest prosodic structure that can be formed with bisyllabic prosodic feet, then they should choose the 2-2 form the most often.

An example of the two types of stimuli is the following:

德文學家: 2-2 versus 1-(2-1)

de2 wen2 xue2 jia1 (Mandarin)

dak1 man4 hok6 gaa1 (Cantonese)

1-(2-1) meaning:

德[[文學]家]

German [[literature] expert]

“A literary scholar of German nationality”

2-2 meaning:

[德文][學家]

[German-language] [learned-scholar]

“An expert of the German language”

剪短一點: 2-2 versus 1-3

jian3 duan3 yi1 dian3 (Mandarin)

zin2 dyun2 jat1 dim2 (Cantonese)

2-2 meaning:

[剪短][一點]

[cut-short] [one-point]

“cut a little”

1-3 meaning:

剪[短一點]

[cut] [short-one-point]

“cut more”

In the first experiment, the main effect we were interested in was the effect of the prosodic foot. Our idea is that if Mandarin has a stronger bisyllabic foot than Cantonese, then we would expect that 2-2 answers would be given more often than 1-(2-1) answers in Mandarin compared to Cantonese. This is also true of the second type of stimuli that we used (2-2 vs. 1-3), where something must be done to incorporate the initial monosyllabic constituent. We should note that it is possible that people could use intonational breaks to form bisyllabic prosodic feet, or perhaps even bisyllabic feet with empty spaces that are not necessarily perceivable (particularly if a space is created at either end of the four-character sequences), and thus use only bisyllabic feet on all types of stimuli. Thus, for instance, the 1-(2-1) stimuli might be read in a *1-break*, *2, 1-break* form, allowing only bisyllabic prosodic feet to be used. However, based on the idea of Duanmu (2002) that taking spaces for prosodic feet is less preferred than filling prosodic feet with actual syllables, and based on creating the simplest prosodic structure, the 2-2 forms should be preferred.

It should be noted that there are, in fact, a number of suggestions about how syllable grouping occurs in Chinese (e.g., Chen, 2000; Duanmu, 2002; Feng, 2002; Shih, 1986). Although the details of the different theories differ, there is common agreement that in Mandarin, bisyllabic feet are generally preferable to any other type (such as trisyllabic or “superfeet”). Therefore, in Mandarin, as far as we can tell, all theories predict that the 2-2 pattern should be prosodically preferable for our stimuli.

Our hypothesis is therefore as follows: if prosody affects reading, it should influence the way people group syllables in Mandarin and Cantonese differently. In particular, we would expect people to most commonly use the 2-2 pattern, and this preference should be stronger in Mandarin compared to Cantonese. Alternatively, if prosody does not affect reading, then we would expect that the two 2-2 pattern would be chosen a similar number of times in Cantonese and Mandarin, because nonprosodic aspects of the stimuli are very similar across the two languages.

Because we do not wish to enter the debate about what are and what are not words in Cantonese and Mandarin, we chose the stimuli based on the individual characters and made sure that the sequences of characters were all typical of Mandarin and Cantonese. (Thus, we did not use sequences like noun–noun–verb–adjective.) We also made sure that the same syntactic bracketing could be used in both languages. The second of these was not very difficult, because, in general, the characters used were all cross-language cognates (or very close to being cognates). One problem we had, however, was that as far as we know, there is no frequency dictionary for Cantonese characters and words (unlike Mandarin). This means that because people do not agree on what words actually are in Cantonese and Mandarin, if word frequencies are generated, there are always a number of implicit assumptions involved. It is thus quite difficult to get participants to try to judge what word and character frequencies are, without getting confounded responses. We therefore had to base frequency comparisons across languages (i.e., rejecting characters that might have large frequency differences) on the objective judgments of the second author who is bilingual, and a number of his informants.

Two types of task were used to examine people’s preferences. In one task, the stimuli were presented quickly and required a speeded decision. In the other task,

the stimuli were presented for as long as necessary, and time pressure was not strictly enforced, although because of the relatively dull nature of performing the experiment, participants did not typically “stop and think” before responding. The idea of using two different types of task was to give some idea about the time course of hierarchical grouping and prosodic processing. If prosodic effects only occur late in the time course of processing, and if our task successfully distinguishes between early and late processing, we would expect an interaction between the two tasks. In this case, because the hierarchical structure of the stimuli are identical in Cantonese and Mandarin, we would expect a very similar pattern of responses in the speeded task, but prosodic effects to emerge in the nonspeeded condition.

EXPERIMENT 1

Participants

Forty-eight students at the University of Hong Kong participated in the experiment. Twenty-eight spoke Cantonese as their first language and 20 spoke Mandarin as their first language. A further 20 students participated in a rating task.

Stimuli

Thirty-one critical stimuli, all with 4 characters, were selected. Sixteen could potentially be read as a 2-2 or 1-(2-1) word structure. Fifteen could potentially be read with a 2-2 or 1-3 structure. None of the words in the sentences had highly frequent uses that were specific to either Cantonese or Mandarin. Two short sentences were constructed for each stimulus, one synonymous with the first meaning and the other synonymous with the second. All of those stimuli appear in Appendix A. A further 100 sentences that varied in length from 3 to 7 characters (20 in each group) and had only one grammatically reasonable structure were used as fillers. Two short sentences were also constructed for each of these sentences, but only one had a meaning synonymous with the sentence meaning.

Two different versions of all of the stimuli were created. One used classic Chinese characters (for the Cantonese speakers), and one used simplified Chinese characters (for the Mandarin speakers). Simplified characters differ from classic characters along various perceptual dimensions, although many are identical. Differences can range from entire characters being different, which is uncommon, to very small changes in the stroke pattern of a character, which is quite common. The underlying morphemes that the characters refer to are still the same, however. A similar phenomenon occurs between some words spelled in British and American English (e.g., *color* and *colour*), although the historical reasons for its occurrence are different. The use of simplified and classic characters was necessary because our Cantonese speakers came from Hong Kong, where classic characters are typically used, and our Mandarin speakers came from areas of Mainland China, where simplified characters are typically used. Apart from individual character differences, the stimuli presented to the Mandarin and Cantonese speakers were identical.

Procedure

An initial semantic congruency check was performed for each of the possible meanings within each stimulus to make sure they were balanced on that dimension. Judges in the task were explicitly informed of the nature of the stimuli and how they should read them. None of the judges were trained linguists. Stimuli were presented next to the meanings. Judges wrote a number from 1 to 7 based on their opinion of the semantic congruency value of each stimuli. Three values were given on the scale: 7 = *reasonable*, 1 = *semantically implausible*, and 3.5 = *semantically unlikely*. We also checked to make sure that the individual words existed or were not extremely rare by using a Mandarin word form frequency dictionary. Unfortunately, as far as we are aware, there is no equivalent Cantonese word form frequency dictionary. However, we did note that none of the Cantonese judges made any queries about nonextant words in the initial judgments, thus suggesting that all the words could be used. The judges also performed semantic congruency judgments for the next experiment at the same time.

The stimuli for the actual task were pseudorandomized for each of the participants. Ten practice trials were used before the test stimuli appeared. For the speeded task, participants were told verbally and via an information screen on the computer that they would briefly see a sentence appear. Following that, they would see two short sentences with different meanings. They were asked to try to choose the meaning most synonymous with the initial sentence as quickly as possible. If they thought both meanings were synonymous, then they were told to choose the meaning that would be the most likely used for the sentence. The individual trials went as follows. First, a plus sign (+) appeared on the screen for 500 ms. Second, a test sentence appeared on the screen for 2000 ms. This was replaced by two short sentences that had different meanings. The two different meanings then remained on the screen until a response was made. For the nonspeeded task, participants were first told that they would see three sentences. They were then told that one of the sentences (the lower sentence) would appear below the other two (the upper sentences), and that the lower sentence was synonymous with one of the upper ones. They were asked to judge which of the upper sentences was synonymous with the lower sentence. If they thought both of the upper sentences were synonymous with the lower sentence, they were told to choose the meaning that they thought would be the best synonym for the sentence. They were asked to do this as accurately as possible. The individual trials went as follows. First, a + appeared in the middle of the screen for 500 ms. Second, the test stimuli (the upper and lower sentences) appeared centered in the middle of the screen. The order of the two upper sentences was counterbalanced across two groups. The sentences disappeared once participants made a response.

Results

To make sure the initial data set was not confounded on semantic congruency, we first compared the average semantic congruency score of the two potential meanings in both stimuli groups. The results showed that the mean semantic congruency values given in the 2-2/1-3 and 2-2/1-(2-1) groups were very

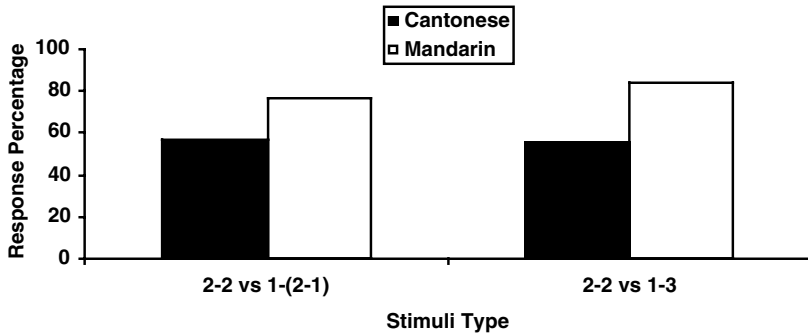


Figure 1. The percentage of responses congruent with the 2-2 meaning in Experiment 1 as a function of stimuli type and language.

similar (2-2/1-3: 6.45/5.87; 2-2/1-[2-1]: 6.31/5.94). A *t* test on the items was not significant for the 2-2/1-(2-1) group, $t(15) = 1.06$, $p = ns$, although another *t* test surprisingly was for the 2-2/1-3 group, $t(14) = 2.28$, $p < .05$. Because the difference in the average was so small in this second group, we ignored it in further analysis. We should note that if such a small difference has an effect on the results, then it should show up as an interaction between the values found in the 2-2/1-3 and 2-2/1-(2-1) groups. The analysis below showed that this did not occur.

In terms of the task results, one item was removed from the analysis completely as it was an idiom in Mandarin. All responses below 100 ms or above 5000 ms were removed from the data set (3.75%). For each participant, all responses 3 *SD* above or below the individual grand mean were considered outliers and removed from the data set (0.63%). Response probabilities were calculated for each participant by dividing the number of times a 2-2 structure was given by the total number of responses minus the number of items removed. An initial examination of the data showed that the two different presentation methods produced very similar results, in terms of the percentage of times that participants gave 2-2 answers. Because this null effect was not easily interpretable based on the data we collected, we examined the results of the two tasks collapsed. The mean results appear in Figure 1.

A 2 (Language) \times 2 (Stimuli Type) analysis of variance examining the probability that participants gave the 2-2 meanings was used to examine the main data set. The results showed that there was a main effect of language, $F_1(1, 46) = 65.26$, $p < .001$; $F_2(1, 28) = 16.83$, $p < .001$, with Mandarin speaking subjects preferring the 2-2 meaning far more often than the Cantonese speaking subjects (80.25 vs. 56.30%). There was no significant effect of stimuli type nor any interactions (all p s $> .08$). To examine whether participants gave more 2-2 responses than would be predicted by chance, a one-group *t* test was performed against a neutral baseline (i.e., 50%). The results showed that the Mandarin speaking participants strongly preferred the 2-2 structure, $t_1(39) = 17.24$, $p < .001$; $t_2(29) = 6.37$, $p < .001$. The Cantonese speaking participants also had a preference for the 2-2 structure, although it was not nearly as strong, $t_1(55) = 3.83$, $p < .001$; $t_2(29) = 1.38$, $p = ns$.

The results we found agree with the arguments of Feng (2002) and others (e.g., Duanmu, 2002), who suggested that the preferred prosodic foot in Mandarin is generally bisyllabic, and that this can be examined via the use of sentences with ambiguous groupings. In particular, Mandarin-speaking participants strongly favored the meaning derived from the 2-2 breakdown of the sentences versus the two alternative structures examined. The same was true to a much weaker extent with the Cantonese speakers, who only appeared to have a mild bias toward the 2-2 structure.

The comparative results from the Mandarin and Cantonese speakers provide evidence that prosody is generated internally when reading. In particular, in Mandarin, the way people resolved syntactic ambiguity appeared to be more affected by a bisyllabic prosodic foot preference than Cantonese. Compared to results from monolingual comparisons, these results have the additional benefit of ruling out within-group extraneous variables that might have otherwise been difficult to control.

One potential problem³ with the stimuli is that in Mandarin, tone 3 sandhi may change the type of groupings people prefer. In our stimuli, we did not control this factor. However, 16 of the stimuli had a tone 3 syllable in them and 14 did not. When divided into two groups, an almost identical pattern of results was found, with respondents choosing 2-2 proportions of .77 and .81 (*ns*) for the group with tone 3 syllables and the group without, respectively. Thus, it appears unlikely that things to do with tone 3 syllables were affecting the results in a meaningful way.

A second potential problem is to do with colloquial versus literary speech. In particular, some of the sentences we used were inadvertently quite literary sounding. Because of this, there is some possibility that people might have used some form of recitation strategy with the literary examples, which could have led to greater rhythmic effects. Given that it is difficult to provide outright rules for determining whether something is literary or colloquial sounding, we divided the sentences up into literary and colloquial categories based on the opinion of one of the authors (which are marked in Appendix A). This led to 12 Mandarin and 15 Cantonese stimuli being considered literary. Surprisingly, the results showed that participants were less likely to use the 2-2 pattern when it was literary in both Mandarin (87.2 vs. 69.7%), $t(28) = 1.91, p < .07$, and Cantonese (62.1 vs. 50.4%), $t(28) = 1.29, p > .1$, although the results were not significant. Despite the lack of significance, the general pattern of results suggests that when literary sentences are used, people may have a tendency to ignore prosodic patterns found in colloquial speech. Although this is potentially interesting, we will not investigate it further.

EXPERIMENT 2

As we discussed in the introduction, the strength of the prosodic influence on ambiguity resolution might be somewhat determined by the predictability of the prosodic context. The previous experiment provided evidence that Mandarin speakers were much more constrained by a bisyllabic prosodic foot than Cantonese speakers. Given that the prosodic foot in Mandarin appeared to have quite a strong influence on syntactic ambiguity resolution, it may be that other constraints that could affect ambiguity resolution, such as semantic expectations

(e.g., Trueswell, Tanenhaus, & Garnsey, 1994), are less likely to influence people's hierarchical grouping choices compared to other languages where the prosodic foot might not have such a strong influence, such as Cantonese. This might be for two different reasons. First, if predictability causes the foot to be used to a greater extent, then more answers congruent with a 2-2 pattern might be given in Mandarin compared to Cantonese. Second, because of potentially different ratios at which bisyllabic and other types of feet are used in Cantonese and Mandarin, more bisyllabic answers might be given in Mandarin even if the prosodic foot was used to exactly the same extent in hierarchical grouping choice.

Investigating the extent that prosodic differences emerge on top of semantic constraints is important, because if differences exist, then it suggests that languages may be differentially sensitive to different types of constraint. That is, if we assume that there are multiple constraints on people's ambiguity resolution preferences, then the way people use those constraints across different languages may differ.

The reason that semantic effects themselves are of interest is that it has been suggested that prosody might override semantic congruency in Chinese in certain circumstances (Feng, 2002). If such a claim is true, it would be very interesting, because it has often been assumed that the grammar parser people use chooses semantically congruent selections, either through interactivity between semantics when syntactic form is being generated (e.g., Boland, 1997) or through revision of the incongruent form (e.g., Frazier, 1979).

Although prosody might be a strong constraint on hierarchical grouping preferences in Mandarin, there are also arguments that semantics plays a large role. In particular, some have argued that processing in Mandarin Chinese has a very strong semantic basis (Lu, 1997; Ma, 1998; Xing, 1995; Xu, 2000). Thus, based on those ideas, it might be predicted that semantic effects could outweigh prosodic effects. However, in comparison with other Chinese languages, the extent that semantics is favored over prosody might not be as much. This is because the relative extent that semantics is used over prosody when resolving ambiguity might be greater in other Chinese languages that do not have such a predictable prosodic foot as Mandarin, such as Cantonese.

To examine the effect of semantic congruency and prosody together, we used the same methodology as Experiment 1, where four-character sentences with two potential meanings were used: either 2-2 or 1-(2-1). However, instead of choosing ambiguous stimuli that had two semantically congruent meanings, we chose stimuli where the 2-2 meaning was semantically incongruent and the 1-(2-1) meaning was not. This type of stimuli is of interest, because Mandarin speakers in the previous experiment did not prefer that form. Therefore, if Mandarin speakers choose the 1-(2-1) form, it means that they are willing to choose semantically congruent forms over prosodically preferred forms. That is, choosing the semantically congruent form represents a choice that is not the typical prosodic default (i.e., 2-2). Our hypothesis is that if Cantonese allows a greater semantic influence on hierarchical grouping resolution than Mandarin, then this should turn up in the data as a stronger preference for the semantically congruent form than the Mandarin speakers. We should note that our semantically incongruent stimuli are extremely incongruent, and therefore, the semantically congruent versus semantically incongruent manipulation is very strong.

An example of the stimuli used in the task is the following:

有否決權

you3 fou3 jue2 quan2 (Mandarin)

jau5 fau2 kyut3 kyun4 (Cantonese)

1-(2-1) (semantically congruent):

有[[否決]權]

have [[no-decide] right]

“have the right to veto”

2-2 (semantically incongruent):

[有否][決權]

[have-not] [decide-right]

“Is it possible to decide something?”

Participants

Forty-eight students at the University of Hong Kong participated in the experiment. Twenty-eight spoke Cantonese as their first language and 20 spoke Mandarin as their first language. The Cantonese speakers participated in the Cantonese experiment and the Mandarin speakers participated in the Mandarin experiment.

Stimuli

Eighteen critical stimuli that all had four characters were used in the experiment. All could either be read as a 2-2 or a 1-(2-1) structure. Unlike the previous experiment, however, only the 1-(2-1) structure was semantically congruent. All of these stimuli appear in Appendix A. The same 100 fillers used in the previous experiment were used in this experiment. In addition, 24 four-character fillers that had two different meanings depending on whether they were read as a 2-2 or 1-(2-1) structure were used. Unlike the critical stimuli, the 1-(2-1) meaning was incongruent and the 2-2 meaning was congruent. Because of a lack of available stimuli, a number of idioms were used as a four-character multiple-meaning control group.

Procedure

The procedure was identical to Experiment 1.

Results

To make sure that semantic congruency was significantly different in the two groups, we compared their average semantic congruency scores given by the participants. One item was completely removed from the analysis because the expected incongruent form was rated more congruent than the congruent form. The results showed that the mean semantic congruency values in the 2-2 (semantically incongruent) versus 1-(2-1) (semantically congruent) groups were judged

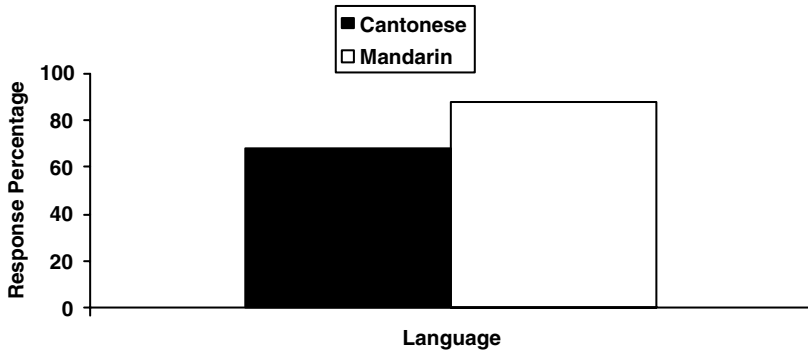


Figure 2. The percentage of semantically congruent 1-(2-1) responses given as a function of language.

to be significantly different at 6.72 versus 1.48, respectively, $t(22) = 29.85$, $p < .001$.

In terms of data processing, the same method was used as the previous experiment. This led to 3.92% of the responses being removed because of having a response time outside the cutoff criterion and 1.10% of the responses were removed because they were outliers. The mean results appear in Figure 2.

The results showed that Mandarin speakers gave more semantically congruent responses, that is, the 1-(2-1) form, than Cantonese speakers (87.79 vs. 67.75%), $t_1(46) = 7.03$, $p < .001$; $t_2(16) = 3.46$, $p < .005$, and both groups preferred the semantically congruent answers more than chance, Mandarin: $t_1(19) = 21.43$, $p < .001$; $t_2(16) = 12.28$, $p < .001$; Cantonese: $t_1(27) = 6.43$, $p < .001$; $t_2(18) = 3.58$, $p < .005$. It is very unlikely that the first of these results was simply due to Cantonese speakers making more errors than Mandarin speakers. This is because there was almost no difference between the Cantonese and Mandarin speaking groups in terms of the proportion of times they chose the 2-2 form in the control group where the 1-(2-1) meaning was semantically incongruent (Cantonese: 81.24%, Mandarin: 83.24%, $t_s < 1$). A post hoc t test examining the difference between the number of semantically congruent responses given in Cantonese compared to the number given in the control group was also significant, $t_1(54) = 4.99$, $p < .001$; $t_2(38) = 2.11$, $p < .05$.

Compared to the last experiment, the results are interesting for two reasons. First, they show that semantic congruency is dominant over prosody in both Cantonese and Mandarin, at least when extremely strong semantic manipulations are made. Second, they show that this is almost completely true of Mandarin, although there still appeared to be some effect of prosody in Cantonese. We infer the second of these observations from the data that suggests Cantonese speakers preferred the semantically congruent 1-(2-1) answers slightly less often than their Mandarin-speaking counterparts, and also less often than the proportion of semantically congruent 2-2 answers they gave in the control group. This second result is somewhat surprising, and goes against our initial predictions that Mandarin

speakers might prefer the 2-2 group more often than Cantonese speakers. Why this result occurred is currently unclear to us.

GENERAL DISCUSSION

Most theories of prosody suggest that there is a constituent, the prosodic foot, that organizes syllables into groups of one or more (see Shattuck-Hufnagel & Turk, 1996, for a review of prosodic hierarchies). The way that the prosodic foot organizes syllables into groups is thought to differ across languages (e.g., McCarthy & Prince, 1993). In this study, we examined the effect of the prosodic foot on hierarchical grouping in Cantonese and Mandarin, two languages that are quite similar grammatically but phonologically very different. The idea was to examine the extent that prosody causes grammatical preference differences in normal reading.

In Experiment 1, we examined the extent that a bisyllabic prosodic foot is used in Cantonese and Mandarin by comparing four-character sentences that were grammatically ambiguous. The ambiguity was such that they could be read two different semantically neutral ways based on the way they were broken down. One of the interpretations was based on a 2-2 reading whereas the other was based on either a 1-(2-1) or 1-3 reading. We hypothesized that, if people have a preference for a bisyllabic prosodic foot, they should prefer the first type of reading, because the other form would need to be read with either a monosyllabic prosodic foot or a more complex strategy that uses either intonational spacing to represent single syllables in a bisyllabic prosodic foot or a bisyllabic foot that breaks word boundaries. The results of the experiment confirmed our expectations that the bisyllabic prosodic foot is more dominant in Mandarin than Cantonese speakers, with Mandarin speakers giving many more 2-2 answers than Cantonese speakers. However, even Cantonese speakers appeared to have a weak preference for the 2-2 form.

In Experiment 2, we investigated the claim that the prosodic foot as a source of constraint may be stronger than semantic congruency in Mandarin (Feng, 2002). If true, such results would be different to the belief that grammatical revision based on semantic variables occurs after an initial syntactic parse. They would also differ to the claim that processing in Mandarin is very semantically driven (Lu, 1997; Ma, 1998; Xing, 1995; Xu, 2000). We investigated this by getting people to judge sentences that could again be read with a 2-2 or 1-(2-1) reading, but where one of the meanings was particularly semantically incongruent and the other was not. The results showed that our Mandarin speakers almost always preferred the semantically plausible structure, even when it meant that they had to choose the structure that was not preferred in the first experiment, that is, the 1-(2-1) structure. Hence, it appeared that semantic congruency was dominant over prosodic typicality. A similar result was found in Cantonese, although it was not as strong. When Cantonese speakers encountered a 2-2 structure that was semantically incongruent, they were more likely to choose it than the Mandarin speakers. The difference between the two groups was quite small, however. Thus, semantics again appeared to dominate prosody, but not completely.

To some extent, the results in Experiment 2, where the Cantonese speakers were affected by prosody more than Mandarin speakers, seems counterintuitive. This is because Mandarin speakers had a much stronger prosodic preference in Experiment 1 when semantic factors were controlled. Thus, at first glance, it appears that if anyone was to be affected by prosody, it should have been the Mandarin speakers. In addition, if a statistical learning system for the constraints, such as that described by Boersma and Hayes (2001), was applied with only information regarding semantics (which we assume to be the same between the two languages) and the typical prosodic foot used (which we assume to differ in at least the extent that it affects final grammatical choices), then it would also not predict the pattern of results found. In this case, because the prosodic foot is probably more predictable (or at least more dominant) in Mandarin than Cantonese, a larger effect of a bisyllabic prosodic foot should have been found in Mandarin, even if it is assumed that the prosodic foot is always used automatically. Thus, what is needed is some explanation as to why prosody might be used more in hierarchical grouping choices in Cantonese compared to Mandarin in semantically nonneutral conditions. At present, however, we only have very speculative reasons for this, and thus we think that this issue needs to be left for further investigation.

Despite the surprising nature of the prosodic influence in Cantonese compared to Mandarin on semantically incongruent items, the results do allow for three firmer and more important conclusions to be drawn. The first relates to the hypothesis of Fodor (1998) that prosody is used in reading and might affect syntactic selection. Our results confirm this (or, more accurately, confirm that there is an effect of prosody on hierarchical grouping choice) by suggesting that at least for Cantonese and Mandarin, there appear to be quite large differences in prosodic effects, even though it was possible to control for potentially confounding variables by using extremely similar items in the two languages. These results support others that have suggested that there are effects of prosodic constituents further up the prosodic hierarchy by showing that a lower level constituent, on which the higher level constituents are dependent, also affects the reading process.

The second conclusion that can be drawn is that both Cantonese and Mandarin speakers appear to be biased toward using a bisyllabic prosodic foot, with Cantonese speakers being less biased than Mandarin speakers. That result agrees with both Yip (1993), who claimed that Cantonese speakers prefer to use bisyllabic prosodic feet in some circumstances, and Feng (2002), who made the same claim about Mandarin speakers. In addition, the result adds further information as to the relative strength of the prosodic foot in the two languages by suggesting that the relative dominance of the bisyllabic prosodic foot is greater in Mandarin than Cantonese.

A third and final conclusion that we will make is that the results of the second experiment suggest that Feng's (2002) claim that prosodic typicality may override semantic ambiguity in hierarchical grouping selection in Mandarin is not general. We found no evidence for it across a large number of stimuli, although we did find strong evidence for a prosodic bias with semantically neutral stimuli. Thus, although prosody may override semantic congruency in relatively specific

circumstances, at least for the stimuli we used, which had extremely incongruent semantic meanings, it did not.

CONCLUSION

In summary, this study examined the effect of the prosodic foot in reading using materials that were otherwise comparable in two different languages. The results we found showed that a difference in the dominance of different types of prosodic foot caused differences in people's hierarchical grouping choices (syntactic selection), even when reading. Furthermore, there appeared to be subtle differences in the interaction between hierarchical grouping, semantics, and prosody, with slightly different effects of these variables found in Cantonese and Mandarin. These results confirm the importance of investigating the role of prosody in silent reading tasks.

APPENDIX A

Mandarin Romanization is in pinyin. Cantonese Romanization is in Jyutping, a form developed by the Linguistics Society of Hong Kong (see <http://cpct92.cityu.edu.hk/lshk/>). Stimuli are presented in the following order: (a) Chinese characters, (b) pinyin (Mandarin) or Jyutping (Cantonese), and (c) a morphosyntactic-group to morphosyntactic-group translation. Responses are presented in the following order: (a) Chinese characters for the two potential responses, (b) the morphosyntactic pattern for the given stimuli, (c) a morphosyntactic-group to morphosyntactic-group translation of the stimuli, and (d) a free translation.

Stimuli were chosen by the second author, but translated by the fourth author. We should note that the opinions of the two authors on the morphosyntactic categories and morphosyntactic bracketing were not in 100% agreement (although they were very similar), which is not surprising because there are many ways to define words in Chinese, and some syntactic categories are not easily categorically distinguished compared to languages like English (e.g., Chao, 1968; Duanmu, 2002; Feng, 2002; Li & Thompson, 1981; Matthews & Yip, 1994). Although it made little difference for the statistical analysis, we used the second author's opinion. However, so that discrepancies can be noted, we present the fourth author's translations and suggested morphosyntactic bracketing in Tables A.1 and A.2.

Table A.1. Stimuli used in Experiment 1 and 2-2 response proportions

Stimuli Mandarin	Stimuli Cantonese	Mandarin		Cantonese		2-2 Response Proportion	
		Response 1	Response 2	Response 1	Response 2	Mandarin	Cantonese
2-2 versus 1-3 彩照相机	彩照相機	1-3 Response 彩色外壳	2-2 Response 彩色相片	1-3 Response 彩色外壳	2-2 Response 彩色相片	0.9	0.78
cai3 zhao4 xiang4 ji1	*coi2 ziu3 soeng2 gei1	A [V-N-N] Colored [take- photo-machine] A camera with a colored outer shell	[A-N] [N-N] [Colored-photo] [photo-machine] A camera that can take colored photos	A [V-N-N] Colored [take- photo-machine] A camera with a colored outer shell	[A-N] [N-N] [Colored-photo] [photo-machine] A camera that can take colored photos		
大红萍果	大紅萍果	很大个的	深色的	很大个的	深色的	0.95	0.07
da4 hong2 ping2guo3	daai6 hung4 ping4gwo2	A [A-N-N] Big [red-apple] A very big apple	[A-A] [N-N] [Very-big] [apple] A dark apple	A [A-N-N] Big [red-apple] A very big apple	[A-A] [N-N] [Very-big] [apple] A dark apple		
好不透明	*好不透明	看不穿	很通透	看不穿	很通透	0.95	0.59
hao4 bu4 tou4ming2	hou2 bat1 tau3 ming4	Adv [Neg-A-A] Very [not- transparent] Very untransparent	[Adv-Neg] [A-A] [Very-not] [transparent] Very transparent	Adv [Neg-A-A] Very [not transparent] Very untransparent	[Adv-Neg] [A-A] [Very-not] [transparent] Very transparent		
*女性温纯	*女性溫純	谈性格	谈女士	谈性格	谈女士	0.89	0.68
nu:3 xing4 wen1chun2	neoi5 sing3 wan1 jau4	N [N-A-A] Female [nature- gentle] This girl is soft and gentle.	[N-N] [A-A] [Female-type] [gentle] The softness of women	N [N-A-A] Female [nature- gentle] This girl is soft and gentle.	[N-N] [A-A] [Female-type] [gentle] The softness of women		

Table A.1 (cont.)

Stimuli Mandarin	Stimuli Cantonese	Mandarin		Cantonese		2-2 Response Proportion	
		Response 1	Response 2	Response 1	Response 2	Mandarin	Cantonese
*响下无声 *xue3 xia4 wu2 sheng1	*雪下無聲 syut3 haa6 mou4 sing1	下雪 N[V-Neg-N] Snow [fall- no-sound] When it's snowing, there is no sound.	在雪下 [N-Post] [Neg-N] [Snow-under] [no-sound] Under the snow, there is no sound.	下雪 N[V-Neg-N] Snow [fall- no-sound] When it's snowing, there is no sound.	在雪下 [N-Post] [Neg-N] [Snow-under] [no-sound] Under the snow, there is no sound.	0.95	0.5
活得精彩 huo2 de2 jing1 cai3	活得精彩 wut6 dak1 zing1 coi2	活着使精彩 [V] [V-N-N] Live [obtain- richness] By being alive, you can do a lot of wonderful things.	生活得很精彩 [V-prt] [Adv-Adv] [Live-able] [rich] Leading a wonderful and fulfilling life	活著使精彩 V [V-N-N] Live [obtain- richness] By being alive, you can do a lot of wonderful things.	生活得很精彩 [V-prt] [Adv-Adv] [Live-able] [rich] Leading a wonderful and fulfilling life	0.89	0.93
剪短一点 jian3 duan3 yi1dian3	*剪短一點 zin2 dyun2 jat1 dim2	剪得短一点 V [A-Num-Cl] Cut [short-one- point] Cut it a bit shorter.	剪了点儿 [V-adv] [Num-CL] [Cut-short] [one-point] Cut short a little bit	剪得短一點 V [A-Num-Cl] Cut [short- one-point] Cut it a bit shorter.	剪了點兒 [V-adv] [Num-CL] [Cut-short] [one-point] Cut short a little bit	0.5	0.19
咬死老鼠 yao3 si3 lao3 shu3	咬死老鼠 au5 sei2 lou5 syu2	老鼠已死 V [A-N-N] Bite [dead-mouse] Bite the dead mouse	杀掉老鼠 [V-V] [N-N] [Bite-die] [mouse] [Bite-die] [mouse] Bit the mouse to its death	老鼠已死 V [A-N-N] Bite [dead-mouse] Bite the dead mouse.	殺掉老鼠 [V-V] [N-N] [Bite-die] [mouse] [Bite-die] [mouse] Bit the mouse to its death	0.89	0.75

丟破花瓶 diu1 po4 hua1 ping2	*丟破花瓶 diu6 po3 faa1 ping4	花瓶已破 V [A-N-N] Drop [broken-vase] Drop the broken vase.	被摔破 [V-V] [N-N] [Drop-break] [vase] Dropped and broke the vase	花瓶已破 V [A-N-N] Drop [broken- vase] Drop the broken vase.	被摔破 [V-V] [N-N] [Drop-break] [vase] Dropped and broke the vase	0.79	0.5
*氣急敗壞 qi4 ji2 bai4 huai4	*氣急敗壞 hei3 gap1 baai6 waai6	‘氣急降 N [A-A-A] Situation [quick- defeat-bad] The situation is suddenly turning against (some entity)	上‘氣不接下‘氣,很狼狽 [N-A] [A-A] [Air-quick] [defeat-bad] Panting a lot; can’t catch up with one’s breath	氣勢急降 N [A-A-A] Situation [quick- defeat-bad] The situation is suddenly turning against (some entity)	上氣不接下氣,很狼狽 [N-A] [A-A] [air-quick] [defeat-bad] Panting a lot; can’t catch up with one’s breath	0.16	0.78
*反經行權 fan3 jing1 hang2 quan2	*反經行權 faan2 ging1 hang4 kyun4	反对这常行的权利 V [Adv-V-N] Object [often- exercise-rights] Object to the rights that are often exercised	行非常用的权利 [Neg-adv] [V-N] [Not-often] [exercise-rights] Exercise rights that are normally not used	反對這常行的權利 V [Adv-V-N] Object [often- exercise-rights] Object to the rights that are often exercised	行非常用的權利 [Neg-adv] [V-N] [Not-often] [exercise-rights] Exercise rights that are normally not used	0.79	0.41
見鬼怕光 jian4 gui3 pa4 guang1	見鬼怕光 gin3 gwai2 paa3 gwong1	鬼愛黑夜 V [N-V-N] See [devil-fear- light] Seen a ghost that is afraid of light	見鬼后,连白天都惊 [V-N] [V-N] [See-devil] [fear-light] After seeing a ghost, (I) have become even afraid of daylight.	鬼愛黑夜 V [N-V-N] See[devil- fear-light] Seen a ghost that is afraid of light	見鬼後,連白天都驚 [V-N] [V-N] [See-devil] [fear-light] After seeing a ghost, (I) have become even afraid of daylight.	1	0.82

Table A.1 (cont.)

Stimuli Mandarin	Stimuli Cantonese	Mandarin		Cantonese		2-2 Response Proportion	
		Response 1	Response 2	Response 1	Response 2	Mandarin	Cantonese
*古往今來 gu3 wang3 jin1 lai2	*古往今來 gu2 wong5 gam1 loi4	由古時到現在 Adv [V-Adv-V] Past [go-now- come] From the ancient time to the present	日子過了,今天已到 [Adv-V] [Adv-V] [Past-go] [now-come] The past's gone; a new day has come.	由古時到現在 Adv [V-Adv-V] Past [go-now- come] From the ancient time to the present	舊日過了,今天已到 [Adv-V] [Adv-V] [Past-go] [now-come] The past's gone; a new day has come.	0.9	0.07
吃素食羊 chi1 su4 shi2 yang2	*吃素食羊 hek3 so3 sik6 joeng4	吃草羊 V [N-V-N] Eat [vegetable-eat- sheep] Eating a vegetarian sheep	吃菜和羊 [V-N] [V-N] [Eat-vegetable] [eat-sheep] Eat vegetables and eat sheep.	吃草羊 V [N-V-N] Eat [vegetable-eat- sheep] Eating a vegetarian sheep	吃菜和羊 [V-N] [V-N] [Eat-vegetable] [eat-sheep] Eat vegetables and eat sheep.	1	0.54
*笑鸡骂狗 xiao4 ji1 ma4 gou3	*笑雞罵狗 siu3 gai1 maa6 gaau2	笑这鸡在骂那狗 V [N-V-N] Laugh [chicken- yell-dog] Laughing at the fact that the chicken was yelling at the dog	嘲笑这鸡, 骂那狗 [V-N] [V-N] [Laugh-chicken] [yell-dog] Laughing at this chicken, while yelling at that dog	笑这雞在罵那狗 V [N-V-N] Laugh [chicken- yell-dog] Laughing at the fact that the chicken was yelling at the dog	嘲笑這雞, 罵那狗 [V-N] [V-N] [Laugh-chicken] [yell-dog] Laughing at this chicken, while yelling at that dog	1	0.76
2-2 versus 1-2-1 是日本人 shi4 ri4 ben3 ren2	*是日本人 si6 jat6 byun2 jan4	1-2-1 Response 是一个日本人 V [[N-N] N] Be [[Japan] person] (Someone) is a Japanese.	2-2 Response 今天, 我... [Adv-N] [Reflexive-N] [Now-day] [self-person] Today, I (formal)	1-2-1 Response 是一個日本人 V [[N-N] N] Be [[Japan]] person] (Someone) is a Japanese.	2-2 Response 今天, 我... [Adv-N] [Reflexive-N] [Now-day] [self-person] Today, I (formal)	0.95	0.11

中医学院 zhong1 yi1 xue2 yuan4	中醫學院 zung1 ji1 hok6 jyun2	在中国的医学院 N [[N-N] N] China [[medicine] school] The medical schools in China	专研中医的地方 [N-N] [A-N] [Chinese-medicine] [learn-institute] Institute of Chinese medicine	中國的醫學院 N [[N-N] N] China [[medicine] school] The medical schools in China	專研中醫的地方 [N-N] [A-N] [Chinese-medicine] [learn-institute] Institute of Chinese medicine	0.95	0.86
*新婚礼法 xin1 hun1 li3 fa3	*新婚禮法 san1 fan1 lai5 faat3	新的婚礼法则 A [[N-N] N] New [[wedding- ceremony] law] A new regulation regarding wedding	新婚时的礼仪 [A-N] [N-N] [New-wedding] [custom-method] The customs and practice of a new wedding	新的禮法則 A [[N-N] N] New [[wedding- ceremony] law] A new regulation regarding wedding	新婚禮的禮儀 [A-N] [N-N] [New-wedding] [custom-method] The customs and practice of a new wedding	0.63	0.6
重工业 zhong4 gong1 ye4 shang1 ye4	重工業 zung6 gung1 soeng1 jip6	着重工商发展 V [[N-N] N] Emphasize [[industry- commerce] trade] Put strong emphasis on industry and commerce	总工业的贸易 [A-N] [N-N] [Total-industry] [trade] The trade of all industries	著重工商發展 V [[N-N] N] Emphasize [[industry- commerce] trade] Put strong emphasis on industry and commerce	總工業的貿易 [A-N] [N-N] [Total-industry] [trade] The trade of all industries	0.89	0.37
德文学家 de2 wen2 xue2 jia1	德文學家 dak1 man4 hok6 gaa1	德国籍文学者 A [[N-N] N] German [[literature] expert] A German expert in literature	德国语文学者 [A-N] [A-N] [German-language] [learned-expert] An expert in German language	德國籍文學者 A [[N-N] N] German [[literature] expert] A German expert in literature	德國語文學者 [A-N] [A-N] [German-language] [learned-expert] An expert in German language	1	0.72

Table A.1 (cont.)

Stimuli Mandarin	Stimuli Cantonese	Mandarin		Cantonese		2-2 Response Proportion	
		Response 1	Response 2	Response 1	Response 2	Mandarin	Cantonese
古代名詞 gu3 dai4 ming2 ci2	古代名詞 gu2 doi6 ming4 ci4	以前用的代名詞 Adv [[V-N] N] Ancient [[replace- name] noun] Pronouns used in ancient times	以前用的名詞 [Adj-N] [N-N]	以前用的代名詞 Adv [[V-N] N] Ancient [[replace- name] noun] Pronouns used in ancient times	以前用的名詞 [Adj-N] [N-N] [Past-time] [noun] Nouns in ancient times	0.95	0.62
*語誤評說 yu3 wu4 ping2 shuo1	*語誤評說 jyu2 m6 ping4 syut3	對語法誤作評論 N [[Adv-V] V] Language [[wrongly- comment] talk] Comment on language use in a careless way	對語法錯誤的評論 [N-N] [V-V] [Language-error] [comment]	對語法誤作評論 N [[Adv-V] V] Language [[wrongly- comment] talk] Comment on language use in a careless way	對語法錯誤的評論 [N-N] [V-V] [Language-error] [comment] Comment on illegal language use	0.35	0.32
*天怒人怨 tian1 nu4 ren2 yuan4	天怒人怨 tin1 jyun3 jan4 lou6	天憤怒, 人哀怨 [N-A] [N-A] [Heaven-angry] [people- complain] The heaven is angry, and the people are upset.	上位者使人民怨聲載道 N [V-N-V] Ruler [make- people-complain]	天憤怒, 人哀怨 [N-A] [N-A] [Heaven-angry] [people- complain] The heaven is angry, and the people are upset.	上位者使人民怨聲載道 N [V-N-V] Ruler [make- people-complain] The ruler has made his people suffer.	0.11	0.28

*成人之惡 cheng2 ren2 zhi1 wu4	*成人之惡 sing4 jan4 zi1 wu3	成年人的丑惡 [A-N] [AM-N] [Grown-man] ['s badness] The ugly side of adults	勉勵並幫助人做壞事 V [N-AM-N] Encourage [man-'s-badness] Encourage others to do bad things.	成年人的醜惡 [A-N] [AM-N] [Grown-man] ['s badness] The ugly side of adults	勉勵並幫助人做壞事 V [N-AM-N] Encourage [man-'s-badness] Encourage others to do bad things.	0.79	0.74
見义勇为 jian4 yi4 yong3 wei4	見義勇為 gin3 ji6 jung5 wai4	行正義勇敢的行为 V [[A-A] N] See [[just-brave] action] Do just and brave things	勇於去做正義的事 [V-N] [Adv-V] [See-justice] [bravely-do] Take the initiative to do just things.	行正義勇敢的行為 V [[A-A] N] See [[just-brave] action] Do just and brave things	勇於去做正義的事 [V-N] [Adv-V] [See-justice] [bravely-do] Take the initiative to do just things.	0.26	0.58
打落水鳥 da2 luo4 shui3 niao3	打落水鳥 daa2 lok6 seoi2 liu5	鳥已落水 V [[V-N] N] Strike [[down- water] bird] Strike the bird which is already in the water.	鳥被击落水 [V-V] [N-N] [Strike-down] [water-bird] The bird was struck into the water.	鳥已落水 V [[V-N] N] Strike [[down- water] bird] Strike the bird which is already in the water.	鳥被擊落水 [V-V] [N-N] [Strike-down] [water-bird] The bird was struck into the water.	0.9	0.56
港大学生: gang3 da4 xue2 sheng5	港大學生 gong2 daai6 hok6 sang1	香港的大学生: N [[A-N] N] Hong Kong [[university] student] Students from universities in Hong Kong	香港大学的学生: [N-N] [A-N] [University of Hong Kong] [student] Students of the University of Hong Kong	香港的大學生 N [[A-N] N] Hong Kong [[university] student] Students from universities in Hong Kong	香港大學的學生 [N-N] [A-N] [University of Hong Kong] [student] Students of the University of Hong Kong	1	0.82

Table A.1 (cont.)

Stimuli Mandarin	Stimuli Cantonese	Mandarin		Cantonese		2-2 Response Proportion	
		Response 1	Response 2	Response 1	Response 2	Mandarin	Cantonese
*无肺病牛 *wu2 fei4 bing4 niu2	無肺病牛 mou5 fai3 beng6 ngau4	健康的牛 Neg [[N-N] N] No [[lung- disease] cow] The cow which doesn't have pneumonia	没有肺部 [Neg-N] [A-N] [No-lung] [sick-cow] A sick cow which doesn't have lungs	健康的牛 Neg [[N-N] N] No [[lung- disease] cow] The cow which doesn't have pneumonia	沒有肺部 [Neg-N] [A-N] [No-lung] [sick-cow] A sick cow which doesn't have lungs	0.79	0.37
无狗拉猫 *wu2 gou3 la1 maol	無狗拉猫 mou5 gau2 lai1 maau1	狗不会拉猫 Neg [N-V-N] No [dog-pull-cat] There are no dogs which will pull cats.	找猫来顶替 [Neg-N] [V-N] [No-dog] [pull-cat] There are no dogs- use cats instead.	狗不会拉猫 Neg [N-V-N] No [dog-pull-cat] There are no dogs which will pull cats.	找猫來頂替 [Neg-N] [V-N] [No-dog] [pull-cat] There are no dogs—use cats instead.	1	0.64
从中学到 cong2 zhong1 xue2 dao4	從中學到 cung4 zung1 hok6 dou3	中学已过 Prep [A-N] Prep From [middle- school] till From middle school to . . .	过程里的所学 [Pre-N] [V-prt] [From-it] [learned] What I learnt from it	中學已過 Prep [A-N] Prep From [middle- school] till From middle school to . . .	過程裡的所學 [Pre-N] [V-prt] [From-it] [learned] What I learnt from it	0.95	0.93

Table A.2. Stimuli used in Experiment 2 and 2-2 response proportions

Stimuli Mandarin	Stimuli Cantonese	1-2-1 Response	2-2 Response	Cantonese		2-2 Response Proportion	
				Response 1	Response 2	Mandarin	Cantonese
高出生率 gao1 chu1 sheng1 lu:4	高出生率 gou1 ceot1 sang1 leot2	多初生儿 A [[V-V] N] High [[birth] rate] High birth rate	超过预计 [V-V] [N-N] [Exceed-cross] [expectation] A rate that has exceeded expectation	多初生兒 A [[V-V] N] High [[birth] rate] High birth rate	超過預計 [V-V] [N-N] [Exceed-cross] [expectation] A rate that has exceeded expectation	0.05	0.26
后青春期 hou4 qing1 chun1 qi1	后青春期 hau6 cing1 ceon1 kei3	约二十岁 Prep [[A-A] N] Late [[puberty] period] Later period of puberty	发情期 [Pre-A] [A-N] [Late-green] [mating-period] Mating period	约二十歲 Prep [[A-A] N] Late [[puberty] period] Later period of puberty	發情期 [Pre-A] [A-N] [Late-green] [mating-period] Mating period	0.05	0.04
扣安全带 kou4 an1 quan2dai4	扣安全带 kau3 on1 cyun4 daai2	避免受伤 V [[A-A] N] Wear [[safety] belt] Wear the seatbelt	带齐东西 [N-N] [Adv-V] [Buckle-name] [all-bring] Bring all things	避免受傷 V [[A-A] N] Wear [[safety] belt] Wear the seatbelt	帶齊東西 [N-N] [Adv-V] [Buckle-name] [all-bring] Bring all things	0.2	0.08

Table A.2 (cont.)

Stimuli Mandarin	Stimuli Cantonese	1-2-1 Response	2-2 Response	Cantonese		2-2 Response Proportion	
				Response 1	Response 2	Mandarin	Cantonese
患关节炎 huan4 guan1 jie2yan2	患關節炎 waan6 gwan1 zit3 jim4	腰酸背痛 V [[N-N] N] Suffer [[joint] inflammation] Suffering from arthritis	患关节炎发了炎 [A-N] [V-N] [Suffered-joint] [become-inflamed] Inflammation at the wounded joints	腰酸背痛 V [[N-N] N] Suffer [[joint] inflammation] Suffering from arthritis	患關節發了炎 [A-N] [V-N] [Suffered-joint] [become-inflamed] Inflammation at the wounded joints	0.05	0.7
办合作社 ban4 he2 zuo4 she4	辦合作社 baan6 hap6 zok3 se5	一起服务 V [[V-V] N] Conduct [[co-operation] society] Serve together	一起作文 [V-Adv] [V-N] [Conduct-together] [compose-society] Compose an essay together	一起服務 V [[V-V] N] Conduct [[co-operation] society] Serve together	一起作文 [V-Adv] [V-N] [Conduct-together] [compose-society] Compose an essay together	0.05	0.11
古建筑物 gu3 jian4 zhu2 wu4	古建築物 gu2 gin3 zuk1 mat6	旧式楼房 A [[V-V] N] Old [[build] thing] An old-styled building	旧式建筑 [Adv-V] [V-N] [Old-build] [build- thing] An old construction	舊式樓房 A [[V-V] N] Old [[build] thing] An old-styled building	舊式建築物 [Adv-V] [V-N] [Old-build] [build- thing] An old construction	0.25	0.59

贺新生儿 he4 xin1 sheng1 er2	賀新生兒 ho6 san1 sang1 ji4	庆祝添丁 V [[adv-V] N] Congratulate [[new-birth] baby] Congratulations on you having a child	朋友贺新生了一个小孩 [N-N] [V-N] [Name] [give birth-baby] (A name) gave birth to a child	慶祝添丁 V [[adv-V] N] Congratulate [[new-birth] baby] Congratulations on you having a child	朋友賀新生了一個小孩 [N-N] [V-N] [Name] [give birth-baby] (A name) gave birth to a child	0.44	0.21
低发病率 di1 fa1 bing lu:4	低發病率 dai1 faat3 beng6 leot2	药物有效 A [[V-N] N] Low [[develop- disease] rate] A low relapse rate	在低处发放病毒比率 [N-V] [N-N] [Low-emit] [sick-rate] The rate of releasing viruses at a low pace	藥物有效 A [[V-N] N] Low [[develop- disease] rate] A low relapse rate	在低處發放病毒比率 [N-V] [N-N] [Low-emit] [sick-rate] The rate of releasing viruses at a low pace	0.05	0.42
新编辑部 xin1 bian1 ji2 bu4	新編輯部 san1 pin1 cap1 bou6	新成立或重组 A [[V-V] N] New [[edit] office] New editing office	新编写的辑部 [Adv-V] [N-N] [New-revised] [edit-part] Newly edited part	新成立或重組 A [[V-V] N] New [[edit] office] New editing office	新編寫的輯部 [Adv-V] [N-N] [New-revised] [edit-part] Newly edited part	0.1	0.62
买印刷机 mai3 yin4 shua4 ji1	買印刷機 maai5 jan3 caat3 gei1	新添了设备 V [[V-V] N] Buy [[print] machine] Buy a new printer	添置了印和一部刷机 [V-N] [V-N] [Buy-stamp] [brush-machine] Buy a stamp and a brushing machine	新添了設備 V [[V-V] N] Buy [[print] machine] Buy a new printer	添置了印和一部刷機 [V-N] [V-N] [Buy-stamp] [brush-machine] Buy a stamp and a brushing machine	0.05	0.42

Table A.2 (cont.)

Stimuli Mandarin	Stimuli Cantonese	1-2-1 Response	2-2 Response	Cantonese		2-2 Response Proportion	
				Response 1	Response 2	Mandarin	Cantonese
大概念化 tai4 gai4 nian4 hua4	大概念化 tai1 koi3 lim6 faa3	不实在 Adv [[N-N]-suffix] Too [[concept]-ize]	过分概括, 理念趋向 [Adv-A] [N-suffix] [Too-general] [concept_ize]	不实在 Adv [[N-N]-suffix] Too [[concept]-ize]	過分概括, 理念趋向 [Adv-A] [N-suffix] [Too-general] [concept_ize]	0.21	0.35
非公式化 fei1 gong1 shi4 hua4	非公式化 fei1 gung1 sik1 faa3	Too abstract 灵活度大 Neg [[N-N] suffix] Not [[formula]ize]	Too general 并非公有, 特式趋向 [Neg-A] [noun- suffix] [Not_public] [method-ize]	Too abstract 靈活度大 Neg [[N-N] suffix] Not [[formula]ize]	Too general 並非公有, 特式趋向 [Neg-A] [noun- suffix] [Not_public] [method-ize]	0	0.35
欠机动性 qian4 ji1 dong4 xing4	欠機動性 him3 gei1 dung6 sing3	Not formulaic 无效率 V [[N-N] N] Lack [[machine- movement] nature] A lack of efficiency	Not public 无机器活动性能 V [N [N-N]] No [machine [activity- function]] No function of mechanical activity	Not formulaic 無效率 V [[N-N] N] Lack [[machine- movement] nature] A lack of efficiency	Not public 無機器活動性能 V [N [N-N]] No [machine [activity] function]] No function of mechanical activity	0.06	0.35
在学术界 zai4 xue2 shu4 jie4	在學術界 zoi6 hok6 seot6 gai3	于一个专业界别 Prep [[A-N] N] At [[learn-skill] field] In academia	在学习艺术界别 [Asp-V] [N-N] [-Ing-learn] [arts-category] Learn different categories in arts	於一個專業界別 Prep [[A-N] N] At [[learn-skill] field] In academia	在學習藝術界別 [Asp-V] [N-N] [-Ing-learn] [arts-category] Learn different categories in arts	0.1	0.48

当秘书长 dang1 mi4 shu1 zhang3	當秘書長 dong3 bei3 syu1 zeong2	行政要員 V [[N-N] N] Become [[secretary] chef] Become head of secretary	看管书的长官 [V-N] [N-N] [Become-secret] [book-chef] Become head in charge of looking after books	行政要員 V [[N-N] N] Become [[secretary] chef] Become head of secretary	看管書的長官 [V-N] [N-N] [Become-secret] [book-chef] Become head in charge of looking after books	0.05	0.22
当驾驶员 dang1 jia4 shi3 yuan2	當駕駛員 dong3 gaa3 sai2 jyun4	司機 V [[V-V] N] Become [[drive] person] Become a driver	保鏢 [V-N] [V-N] [Become- important person] [execute-person] Become a bodyguard	司機 V [[V-V] N] Become [[drive] person] Become a driver	保鏢 [V-N] [V-N] [Become- important person] [execute-person] Become a bodyguard	0	0
有否决权 you3 fou3 jue2 quan2	有否決權 jau2 fau2 kyut3 kyun4	可以反对 V [[V-V] N] Have [[no-decide] right] Have the right to veto	可以決定嗎? [V-Neg] [V-N] [Have-not] [decide-right] Can it be decided?	可以反對 V [[V-V] N] Have [[no-decide] right] Have the right to veto	可以決定嗎? [V-Neg] [V-N] [Have-not] [decide-right] Can it be decided?	0.35	0.29

Note: Asterisked terms are literary. A, adjective; Adv, adverb; AM, adjective marker; Cl, classifier; N, noun; Neg, negator; Num, number; Prep, preposition; Post, postposition; prt, verbal particle; V, verb.

ACKNOWLEDGMENTS

Conrad Perry was supported by a UDF grant from the University of Hong Kong.

NOTES

1. We use Jyutping as our Romanized script for Cantonese. This system was developed and described by the Linguistic Society of Hong Kong.
2. When people talk about different Chinese languages, they often use the word dialect. This word is used more in a sociolinguistic sense, in that most people who speak Chinese dialects are Chinese, than in a statistical one, where measures such as phonological overlap across languages, shared words, pragmatic usage, and so forth, might be compared (e.g., Cheng, 1997). Cantonese and Mandarin, for example, have very little mutual intelligibility, whereas Spanish and Italian have some, but we never hear people call Spanish and Italian different dialects of the same language. We therefore use the word language rather than dialect in this paper.
3. We are grateful to an anonymous reviewer for pointing out these problems. A third potential problem is that two of the Mandarin stimuli used a second syllable that could have been pronounced using a neutral tone by some speakers. If the stimuli are pronounced in that way the only likely bracketing available to the speakers would have been the 2-2 one.

REFERENCES

- Bates, E., Chen, S., Li, P., Opie, M., & Tzeng, O. (1993). Where is the boundary between compounds and phrases in Chinese? A reply to Zhou et al. *Brain and Language*, 45, 94–107.
- Bates, E., Devescovo, A., & D'Amico, S. (1999). Processing complex sentences: A cross linguistic study. *Language and Cognitive Processes*, 14, 69–123.
- Bauer, R. S., & Benedict, P. K. (1997). *Trends in linguistics. Studies and monographs* (Vol. 102). Berlin: Mouton de Gruyter.
- Boersma, P., & Hayes, B. (2001). Empirical tests of the gradual learning algorithm. *Linguistic Inquiry*, 32, 45–86.
- Boland, J. E. (1997). The relationships between syntactic and semantic processes in sentence comprehension. *Language and Cognitive Processes*, 12, 423–484.
- Chao, Y. R. (1968). *A grammar of spoken Chinese*. Berkeley, CA: University of California Press.
- Chen, M. Y. (2000). *Tone sandhi: Patterns across Chinese dialects*. Cambridge: Cambridge University Press.
- Cheng, C.-C. (1997). Measuring relationship among dialects: DOC and related resources. *Computational Linguistics and Chinese Language Processing*, 2, 41–72.
- Cohen, M. A., & Grossberg, S. (1986). Neural dynamics of speech and language coding: Developmental programs, perceptual grouping, and competition for short-term memory. *Human Neurobiology*, 5, 1–22.
- Cuetos, F., & Mitchell, D. C. (1988). Cross-linguistic differences in parsing: Restrictions on the use of the late closure strategy in Spanish. *Cognition*, 30, 73–105.
- Cuetos, F., Mitchell, D. V., & Corley, M. M. B. (1996). Parsing in different languages. In M. Cerreiras, J. Garcia-Albea, & N. Sebastian-Galles (Eds.), *Language processing in Spanish* (pp. 145–187). Mahwah, NJ: Erlbaum.
- Duanmu, S. (2002). *The phonology of standard Chinese*. New York: Oxford University Press.
- Duanmu, S. (2004). Left-headed feet and phrasal stress in Chinese. *Cahiers de Linguistique-Asie Orientale*, 33, 65–103.
- Feng, S. (2001). Prosodic structure and compound words in Classical Chinese. In J. Packard (Ed.), *New approaches to Chinese word formation: Morphology, phonology and the lexicon in modern and ancient Chinese* (Vol. 1997, pp. 197–260). Berlin: Mouton de Gruyter.

- Feng, S. (2002). *Lincom studies in Asian linguistics: Vol. 44. The prosodic syntax of Chinese*. Munich: Lincom Europa.
- Flynn, C. (2004). *Intonation in Cantonese*. Munich: Lincom.
- Fodor, J. D. (1998). Learning to parse? *Journal of Psycholinguistic Research*, 27, 285–319.
- Fodor, J. D. (2002). Prosodic disambiguation in silent reading. In M. Hirotani (Ed.), *Proceedings of NELS 32*. Amherst, MA: University of Massachusetts, GLSA.
- Frazier, L. (1979). *On comprehending sentences: Syntactic parsing strategies*. Bloomington, IN: Indiana University Linguistics Club.
- Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.
- Lu, J. M. (1997). On semantic directionality analysis [Guanyu yuyi zhixiang fengxi]. *Forum on Chinese Linguistics [Zhongguo Yuyanxue Luncong]*, 1, 34–48.
- Ma, Q. Z. (1998). *Categories in Chinese semantic grammar [Hanyu yuyi yufa fanzhou wenti]*. Beijing: Beijing University of Language and Culture Press.
- Mathews, S., & Yip, V. 1994. *Cantonese: A comprehensive grammar*. London: Routledge.
- McCarthy, J., & Prince, A. (1993). *Prosodic morphology: Constraint interaction and satisfaction* (Tech. Rep. RuCCS-TR-3). New Brunswick, NJ: Rutgers University, Rutgers Center for Cognitive Science.
- Packard, J. L. (2000). *The morphology of Chinese*. Cambridge: Cambridge University Press.
- Selkirk, E. (1980). The role of prosodic categories in English word stress. *Linguistic Inquiry*, 11, 563–605.
- Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, 25, 193–247.
- Shih, C.-L. (1986). *The prosodic domain of tone sandhi in Chinese*. Unpublished doctoral dissertation, University of California, San Diego.
- Smolensky, P. (2000). Grammar-based connectionist approaches to language. *Cognitive Science*, 23, 589–613.
- Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: use of thematic role information in syntactic disambiguation. *Journal of Memory and Language*, 33, 285–318.
- Wong, W. Y. P., Chan, M. K.-M., & Beckman, M. E. (2004). An autosegmental-metrical analysis and prosodic annotation conventions for Cantonese. In A. S. Jun (Ed.), *Prosodic typology—The phonology of intonation and phrasing* (pp. 271–300). New York: Oxford University Press.
- Xing, F. Y. (1995). *The study of Chinese grammar [Hanyu yufaxue]*. Changchun, China: North-East Normal University Press [Dongbei Shifan Daxue Chubanshe].
- Xu, T. Q. (2000). *On language [Yuyan lun]*. Changchun, China: North-East Normal University Press [Dongbei Shifan Daxue Chubanshe].
- Yip, M. (1993). Cantonese loanword phonology and optimality theory. *Journal of East Asian Linguistics*, 2, 261–291.