Phrase Lengths and the Perceived Informativeness of Prosodic Cues in Turkish

Nazik Dinçtopal Deniz
Boğaziçi University, Turkey

Janet Dean Fodor
The City University of New York, USA

Abstract

It is known from previous studies that in many cases (though not all) the prosodic properties of a spoken utterance reflect aspects of its syntactic structure, and also that in many cases (though not all) listeners can benefit from these prosodic cues. A novel contribution to this literature is the Rational Speaker Hypothesis (RSH), proposed by Clifton, Carlson and Frazier. The RSH maintains that listeners are sensitive to possible reasons for why a speaker might introduce a prosodic break: "listeners treat a prosodic boundary as more informative about the syntax when it flanks short constituents than when it flanks longer constituents", because in the latter case the speaker might have been motivated solely by consideration of optimal phrase lengths. This would effectively reduce the cue value of an appropriately placed prosodic boundary. We present additional evidence for the RSH from Turkish, a language typologically different from English. In addition, our study shows for the first time that the RSH also applies to a prosodic break which conflicts with the syntactic structure, reducing its perceived cue strength if it might have been motivated by length considerations. In this case, the RSH effect is beneficial. Finally, the Turkish data show that prosody-based explanations for parsing preferences such as the RSH do not take the place of traditional syntax-sensitive parsing strategies such as Late Closure. The two sources of guidance co-exist; both are used when available.

Keywords

Sentence processing, prosodic phrasing, Turkish, Late Closure, Rational Speaker Hypothesis

Introduction

Since the classic study by Lehiste (1973), a growing number of experiments have demonstrated that the prosody of an utterance can provide listeners with cues to its syntactic structure (e.g.,

Corresponding author:
Nazik Dinçtopal Deniz, Boğaziçi University, Faculty of Education, Department of Foreign Language Education, Bebek, 34342, Istanbul, Turkey.
Email: nazik.dinctopal@boun.edu.tr

Kjelgaard & Speer, 1999; Marslen-Wilson, Tyler, Warren, Grenier, & Lee, 1992; Nagel, Shapiro, Tuller, & Nawy, 1996; Price, Ostendorf, Shattuck-Hufnagel, & Fong, 1991; Schafer, Speer, Warren, & White, 2000; Speer, Kjelgaard, & Dobroth, 1996; Stoyneshka, Fodor, & Fernández, 2010). This increased attention to the prosodic properties of spoken language brings with it the potential for new kinds of explanations for the performance of the human sentence processing mechanism, in addition to or in place of the traditional syntactically defined parsing strategies such as Minimal Attachment and Late Closure in the early Garden Path model (Frazier, 1978). The latter were based largely on data from silent reading in the absence of any overt prosodic information. Here we will consider an explanation for certain syntactic parsing preferences which attributes them to biases induced by prosodic phrase lengths.

Clifton and colleagues (Clifton, Carlson, & Frazier, 2002, 2006) have introduced a novel and interesting prosody-based explanation for interpretive preferences for noun phrase (NP) coordination and adverbial phrase ambiguities. Their Rational Speaker Hypothesis (RSH) predicts that prosodic breaks flanking short constituents are treated by perceivers as more informative about the syntactic structure of an utterance than prosodic breaks flanking long constituents. The rationale is that if a prosodic break might have been produced by the speaker in order to divide up an over-long constituent, then listeners could not confidently rely on it as a valid indicator of syntactic structure; by contrast, a prosodic break that is not required by purely prosodic considerations is more likely to be construed by listeners as motivated by the syntactic structure, and thus could have a greater impact on the structure assigned to the sentence.

A note on terminology: Despite its name, the RSH is not a psycholinguistic hypothesis about the behavior of speakers. It is a hypothesis about the behavior of speakers which, according to Clifton and colleagues, is entertained by listeners. It affects listeners' interpretation of the locations of prosodic boundaries in the speech they hear. If a boundary is not (or could not be) motivated by purely prosodic (metrical) considerations such as phrase length, it is likely to be deemed by a listener to reflect alignment with syntactic structure – on the assumption that a rational speaker would not introduce a prosodic boundary for no cause. In its practical implications for making parsing decisions, however, the RSH is equivalent to a simple parsing strategy for listeners: in computing syntactic tree structure, give more weight to prosodic boundaries that flank short constituents than to boundaries that flank longer constituents (Clifton et al., 2006, p. 855). As such, evaluation of RSH requires data not from production experiments but from perception (listening) experiments, such as Clifton and colleagues have provided and as we will offer here. In short: in what follows, we will discuss the RSH as a parsing strategy for spoken language.

In one experiment Clifton et al. (2006) investigated the RSH with NP coordination ambiguities as in (1) and (2). Sentences had either short NP conjuncts as in (1) or long NP conjuncts as in (2). The sentences were presented auditorily, with intonational phrase (IPh) boundaries at locations indicated here by ||.

(1) Short NP
     a. Pat || or Jay and Lee || convinced the bank president to extend the mortgage.
     b. Pat or Jay || and Lee || convinced the bank president to extend the mortgage.

(2) Long NP
     a. Patricia Jones || or Jacqueline Frazier and Letitia Connolly || convinced the bank
         president to extend the mortgage.
     b. Patricia Jones or Jacqueline Frazier || and Letitia Connolly || convinced the bank
         president to extend the mortgage.

                                                                                    (Clifton et al., 2006, p. 855)

The early IPh boundary in (1a) was predicted to bias the listeners towards the interpretation in which either Pat (one person) or Jay and Lee (two people) convinced the bank president; and likewise in (2a). By contrast, the late prosodic boundary in (1b) was predicted to bias the listeners towards the interpretation in which either Pat or Jay (one of two people) and Lee (one person) convinced the bank president; and likewise in (2b). The phrase length differences in the materials provided a test of the RSH. Participants mostly chose the prosodically appropriate interpretation for both short and long NPs, but did so significantly more often for the short NPs. Thus, as the RSH predicts, these listeners treated a prosodic boundary as more informative about the syntax when it flanked short constituents (as in (1)) than when it flanked long constituents (as in (2)).

No other evidence for RSH has been reported since Clifton et al. (2006; though see Hwang & Schafer, 2009). However, one may look for possible RSH effects in previous experiments by other researchers, which were not designed or discussed with RSH phenomena in mind. For example, a study by Kjelgaard and Speer (1999) found less difficulty for a late closure (LC) structure over an early closure (EC) structure when a prosodic boundary was in a misleading position. Kjelgaard and Speer investigated English LC/EC ambiguities such as in (3). (|| marks prosodic boundary, / marks syntactic boundary and boldface indicates contrastive focus whose effect is to elicit neutral prosody (i.e., no boundary) elsewhere in the sentence; see Speer et al., 1996, pp. 256–257.)

(3)   a. Cooperating prosody, LC syntax: When Roger leaves the house || / it's dark.
      b. Cooperating prosody, EC syntax: When Roger leaves || / the house is dark.
      c. Conflicting prosody, LC syntax: When Roger leaves || the house / it's dark.
      d. Conflicting prosody, EC syntax: When Roger leaves / the house || is dark.
      e. Neutral prosody, LC syntax: When **Roger** leaves the house / it's dark.
      f. Neutral prosody, EC syntax: When **Roger** leaves / the house is dark.

(Kjelgaard & Speer, 1999, p. 156)

Tasks were acceptability judgment, end-of-sentence comprehension, and cross-modal naming. Within their respective cooperating prosody conditions, LC and EC structures were processed equally efficiently. In the neutral prosody conditions, there was an advantage for the LC structure, which could be attributed to the parser following the syntactic processing strategy of Late Closure (Frazier, 1978) in the absence of prosodic cues. The observation of main interest in relation to the RSH is the finding of an advantage for the LC structure in the conflicting prosody conditions. This too might be explained in terms of the parser relying on a traditional syntactic Late Closure strategy (though Kjelgaard and Speer consider several alternative accounts, for example, reanalyzing the intended LC structure as an EC structure containing a topicalized NP in the second clause). But there is also the possibility under the RSH that the misleading prosody was more difficult to ignore in the EC items because it was perceived as more informative there than in the LC items, since the prosodic break in the EC items was flanked by a short constituent (*is dark*).

This raises the possibility that the prosody-sensitive RSH might offer an explanation for various other findings that would traditionally be attributed to the syntactically defined Late Closure strategy. At the extreme, RSH might thereby render the syntactic Late Closure strategy redundant. If that were the case it would be a noteworthy finding, since the Late Closure strategy is still widely accepted, at least for constructions which involve 'primary relations' (syntactic argument structure, as in (3)), even though the Construal Theory of Frazier and Clifton (1996) rejects the Late Closure strategy for 'non-primary relations' such as adverbial and coordinate constructions.

In order to examine the possibility of competition between prosody-based and syntax-based influences on sentence processing, phrase lengths must be tested which would run counter to Late Closure influences, if RSH is correct, as well as phrase lengths which would reinforce Late Closure (as in the Kjelgaard and Speer study). That is the purpose of the present study.

We present an experiment which investigates a possible interplay between RSH length effects and Late Closure. The experimental materials are Turkish sentences. In examining the relation between syntactic phrases and prosodic phrases, Turkish has the advantage that alignment at the lower prosodic level (word level) is very well-behaved. In Turkish every lexical word is realized as a prosodic word (PWd; Inkelas & Orgun, 2003), making it easy to systematically manipulate prosodic phrase lengths at the higher level. Also, Turkish differs typologically from English both syntactically and prosodically, so if RSH does manifest itself in Turkish that could provide insight concerning the generality of RSH effects, offering at least a hint that RSH might be universal.

Our experiment addressed three questions: (i) Does the RSH apply robustly in Turkish? That is: Are Turkish speakers influenced by phrase lengths in evaluating the perceived informativeness of prosodic boundaries? (ii) If so, does this occur both when a boundary aligns with a syntactic break (potentially helpful) and also when it does not (potentially disruptive, as contemplated above for Kjelgaard and Speer's construction (3c, d))? (iii) When RSH effects are pitted against potential Late Closure effects, can they be distinguished from each other and if so, does one or other prevail? In order to answer these questions, we designed a listening experiment which systematically manipulates the lengths of prosodic phrases in LC and EC syntactic structures. If differences in processing difficulty were entirely due to syntactically-driven parsing strategies such as Late Closure, the outcomes of the length conditions would be similar, without any significant interaction between length, prosody and syntax. However, if prosodic phrase lengths play a role in syntactic processing, in the manner that RSH predicts, different outcomes for the different length conditions are expected, as detailed below in the Predictions section.

2   The present study

2.1   Nature of the ambiguity

The Turkish ambiguity we studied concerns the attachment of a phrase as part of a subject NP or standing alone as an object NP. This is illustrated in (4a, b), where the word *psikoloğ-u* has an ambiguous suffix that is compatible with either subject or object status. The ambiguity is resolved by the verb that follows: *sev-il-di* or *sev-diğ-i-ni*.
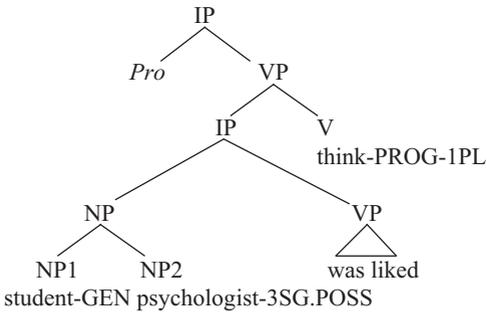
(4)   a. LC: Ø   Öğrenci-nin psikoloğ-u   ‖   sev-il-di   san-ıyor-uz.
              Pro student-GEN psychologist-3SG.POSS-   like-PASS-PAST   think-PROG-1PL
              'We think that the student's psychologist was liked.'

      b. EC: Ø   Öğrenci-nin ‖ psikoloğ-u   sev-diğ-i-ni   san-ıyor-uz.
              Pro student-GEN   psychologist-ACC   like-FN-3SG.POSS-ACC think-PROG-1PL
              'We think that the student liked the psychologist.'

Turkish is a head-final pro-drop language: verbs follow their subject and objects, and the subject may be phonologically null. All target sentences consisted of two clauses, as in (4): a complement clause was followed by a main clause which had a phonologically null subject (acceptable in Turkish even in the absence of a preceding referent for it). The complement clause took two different forms, with different morphology on its verb, but in both cases it contained an overt subject and a verb phrase.

As noted, the temporary ambiguity arises because of the homophony of the morphological marking of the second noun (*psychologist*). In both cases the morpheme is –*u* but in (4a) it marks the noun as 3rd person singular possessive, while in (4b) it is an accusative case marker on the direct object. In the LC structure (4a), the first and second noun (*student* and *psychologist*) form a complex NP, which has no overt nominative case marking but functions as the nominative subject of the passive verb *sev-il-di*. This amounts to late closure of the subject: *student-GEN psychologist-3SG.POSS*. In the EC structure (4b), only the first noun constitutes the subject (a genitive-marked subject). The subject must be closed early, ending at *student-GEN*, because the subordinate verb (*sev-diğ-i-ni* in (4b)) is active and transitive, leading *psikoloğ-u* to be interpreted as its accusative object: *psychologist-ACC*. Other relevant details of these sentences are presented in Dinçtopal-Deniz (2014).

The syntactic structures of these sentences are shown in simplified form in the diagrams in (5), using English glosses. Though not identical to the English structures tested by Kjelgaard and Speer (1999), these Turkish structures likewise involve primary relations, since the ambiguity concerns whether the ambiguous NP is the head of a larger NP or is the object of the embedded verb.

(5a): Late closure of the subject (5b): Early closure of the subject
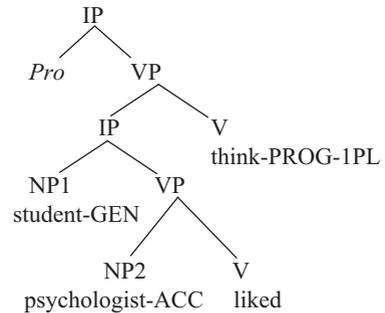
The two constructions differ in their prosodic phrasing, indicated by the position of the prosodic boundary marker || in (4a, b) above. The subject is the default topic in the canonical subject–object–verb word order in Turkish (Erguvanlı, 1984). Topics are followed by an IPh boundary, realized as a sharp rise to a final high boundary tone with a following optional pause (Kamali, 2008; Vallduví & Engdahl, 1996). The location of that IPh boundary provides a clear phonological cue to where the subject ends in (4a) (after *psychologist*) and (4b) (after *student*), eliminating the ambiguity before the disambiguating morphology on the embedded verb becomes available.

2.2 Phrase length manipulation

As a basis for phrase length variations in the experimental materials, all target items consisted of six PWds. Starting with constructions as in (4), an extra word was added to the subject in both LC and EC versions (e.g., *yedi* (seven) in (6) and (7) below), creating a basic five-prosodic-word pattern. Then either the subject or the subordinate verb phrase (VP) was lengthened by addition of another word (e.g., *yaklaşık* (nearly) in (6) and *oldukça* (much) in (7)). The lengthening words did not introduce an additional prosodic boundary or add significantly to the meaning of the sentence.

These items were pronounced with cooperating, conflicting and neutral prosody, as in Kjelgaard and Speer (1999). In the cooperating and conflicting prosody conditions, the 6 PWds were grouped

into two prosodic phrases, but the groupings differed between 2+4 PWds, 3+3 PWds, and 4+2 PWds. This choice of lengths of the material before and after the prosodic break was based on an analysis of pause frequencies by Nash (1973). Nash's data revealed that readers paused at every 4.2 words on average (with a range from 2.9 to 7.8 words). Therefore, prosodic phrases consisting of 2 PWds would be perceived as atypically short, while those with 3 or 4 PWds would fall within the normal range. Hence, only the 3+3 PWds grouping was fully optimal.

2.3   Pre-tests

Before the main experiment, all target sentences underwent several pre-tests which are briefly summarized here; details are provided in Appendix A. Target items were pre-tested for semantic plausibility: a corpus analysis and a normative study were conducted to ensure that none of the target items used in the experiment was inherently biased towards either an LC or an EC interpretation. Then the 24 items selected on the basis of the normative study were read aloud and recorded by the first author (a native speaker of Turkish) for use in the main experiment. The items were pronounced with cooperating, conflicting, and neutral prosody (sample pitch tracks are shown in Figures 1–6 in the Materials section below). These spoken sentences were then pre-tested for their prosodic properties via a pronunciation acceptability judgment task. Pronunciation acceptability ratings confirmed that all the recordings had the intended prosody. This was further confirmed by an acoustic analysis of durations and fundamental frequency (F0) in the temporally ambiguous regions of the target items. See Appendix A for details of pre-tests.

2.4   Method

2.4.1   Design. To test whether the pattern of phrase lengths influences the perceived informativeness of prosodic cues, the target items in the experiment differed systematically in their phrase lengths as well as their prosody and syntactic structure. The complexity of the design made it impractical to test the length contrast within participants, since it would have greatly multiplied the number of items for each participant. The length manipulation was therefore a between-subjects condition, whereas the prosody and syntactic structure were manipulated within subjects.

2.4.2   Participants. The experiment was conducted in Turkey. 106 native speakers of Turkish took part in the experiment. 54 of those participants were in the lengthened subject condition (mean age = 25.2, 17 male) and 52 were in the lengthened VP condition (mean age = 26.2, 14 male). Participants received 15 Turkish liras (~$8.5 at the time of the experiment) for their participation.

2.4.3   Materials. The conditions are illustrated in (6) and (7), which provide English translations for the sentences (same for a., c. and e., and for b., d. and f.). Word glosses are not included here in order to focus on the length manipulation across conditions. (The reader may refer to (4) for glosses for a 4 PWd pattern; translations for the lengthening words are given in the Phrase Length Manipulation section above.) A set of items consisted of 6 versions, differing in syntax (LC, EC) and in prosodic phrase lengths (2+4, 3+3, 4+2). One complete set is shown in (6) for the lengthened subject condition and in (7) for the lengthened VP condition. As above, the symbol ‖ is used to mark prosodic boundaries, / is used to mark syntactic boundaries, boldface indicates contrastive focus. In cooperating prosody, ‖ and / coincide. In conflicting prosody they are at different locations in the sentence. In neutral prosody there is / but no ‖. Also indicated in (6) and (7) are the phrase lengths in PWds of the subject and the VP sequence in each version. Prosodically non-optimal lengths include an unduly short phrase of just two PWds.

(6) Lengthened subject

    a. Cooperating prosody – LC syntax – non-optimal length (4+2 PWds)

      Yaklaşık yedi öğrencinin psikoloğu ‖ / sevildi sanıyoruz.
      'We think that the psychologist of nearly seven students was liked.'

    b. Cooperating prosody – EC syntax – optimal length (3+3 PWds)

      Yaklaşık yedi öğrencinin ‖ / psikoloğu sevdiğini sanıyoruz.
      'We think that nearly seven students liked the psychologist.'

    c. Conflicting prosody – LC syntax – optimal length (3+3 PWds)

      Yaklaşık yedi öğrencinin ‖ psikoloğu / sevildi sanıyoruz.

    d. Conflicting prosody – EC syntax – non-optimal length (4+2 PWds)

      Yaklaşık yedi öğrencinin / psikoloğu ‖ sevdiğini sanıyoruz.

    e. Neutral prosody – LC syntax (6 PWds)

      Yaklaşık **yedi** öğrencinin psikoloğu / sevildi sanıyoruz.

    f. Neutral prosody – EC syntax (6 PWds)

      Yaklaşık **yedi** öğrencinin / psikoloğu sevdiğini sanıyoruz.

(7) Lengthened VP

    a. Cooperating prosody – LC syntax – optimal length (3+3 PWds)

      Yedi öğrencinin psikoloğ-u ‖ / oldukça sevildi sanıyoruz.
      'We think that the psychologist of seven students was much liked.'

    b. Cooperating prosody – EC syntax – non-optimal length (2+4 PWds)

      Yedi öğrencinin ‖ / psikoloğu oldukça sevdiğini sanıyoruz.
      'We think that seven students liked the psychologist much.'

    c. Conflicting prosody – LC syntax – non-optimal length (2+4 PWds)

      Yedi öğrencinin ‖ psikoloğu / oldukça sevildi sanıyoruz.

    d. Conflicting prosody – EC syntax – optimal length (3+3 PWds)

      Yedi öğrencinin / psikoloğu ‖ oldukça sevdiğini sanıyoruz.

    e. Neutral prosody – LC syntax – (6 PWds)

      Yedi öğrencinin psikoloğu / oldukça sevildi sanıyoruz.

    f. Neutral prosody – EC syntax – (6 PWds)

      Yedi öğrencinin / psikoloğu oldukça sevdiğini sanıyoruz.

In cooperating prosody conditions, there was an IPh boundary after the subject, in accord with the topicalization facts in Turkish, noted above. In conflicting prosody conditions, the speaker used EC prosody for LC sentences and LC prosody for EC sentences. This moved the prosodic break earlier for the LC sentences (before *psikoloğu* in the examples above) and later for the EC sentences (after *psikoloğu*). Rather than cross-splicing recordings to create conflicting prosody, which may distort the materials acoustically, the conflicting prosody conditions were spoken. In the neutral prosody conditions, the items had no prosodic boundary anywhere in the sentence, thus were not

prosodically biased toward either syntactic parse; they contained no prosodic cue to the syntactic structure of the sentence. In order for this neutral prosody to sound natural, the modifier (*yedi*) of the subject phrase in both (6) and (7) received a contrastive accent, increasing both its duration and intensity (Ipek, 2011). This tends to reduce prosodic variation in the remainder of the sentence. (See Figures 5 and 6 for details, and Kjelgaard and Speer (1999) for a similar technique.)

The pitch tracks in Figures 1–6 exemplify the three prosody conditions (cooperating, conflicting and neutral), all shown here in the lengthened subject condition only. (See Table A.3.2 in Appendix A.3 for F0 and timing measurements across all experimental items).
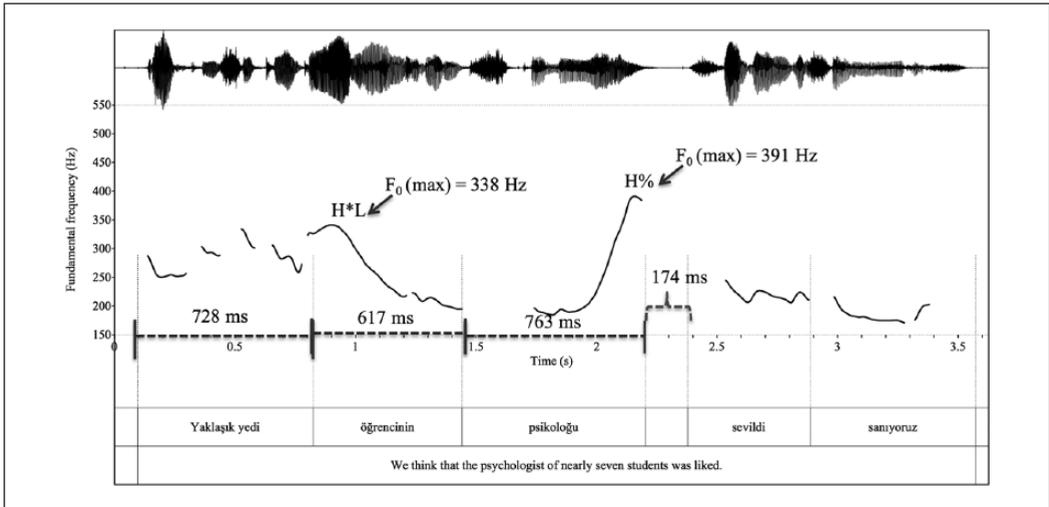

Figure 1. Waveform and pitch track for one LC sentence ((6a) above) uttered with cooperating prosody.
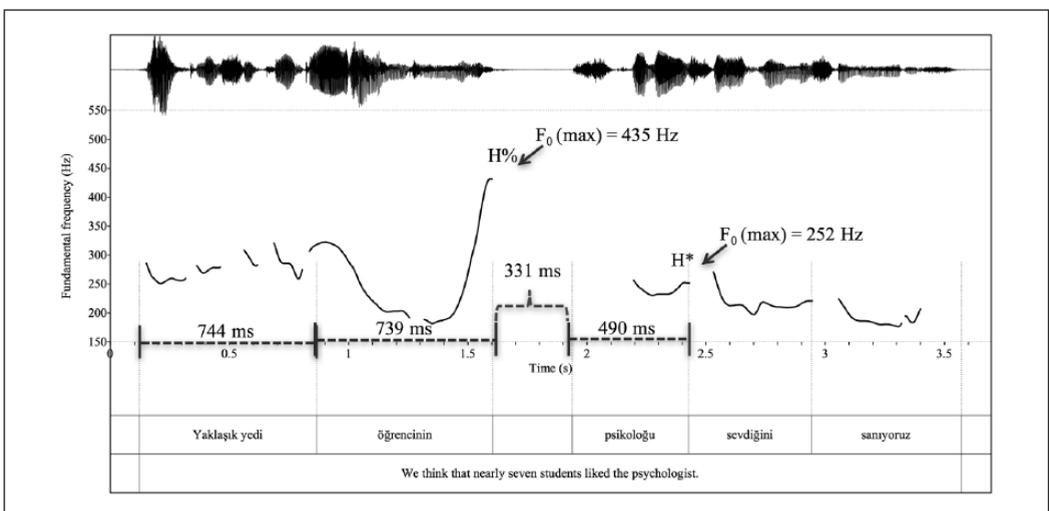

Figure 2. Waveform and pitch track for one EC sentence ((6b) above) uttered with cooperating prosody.
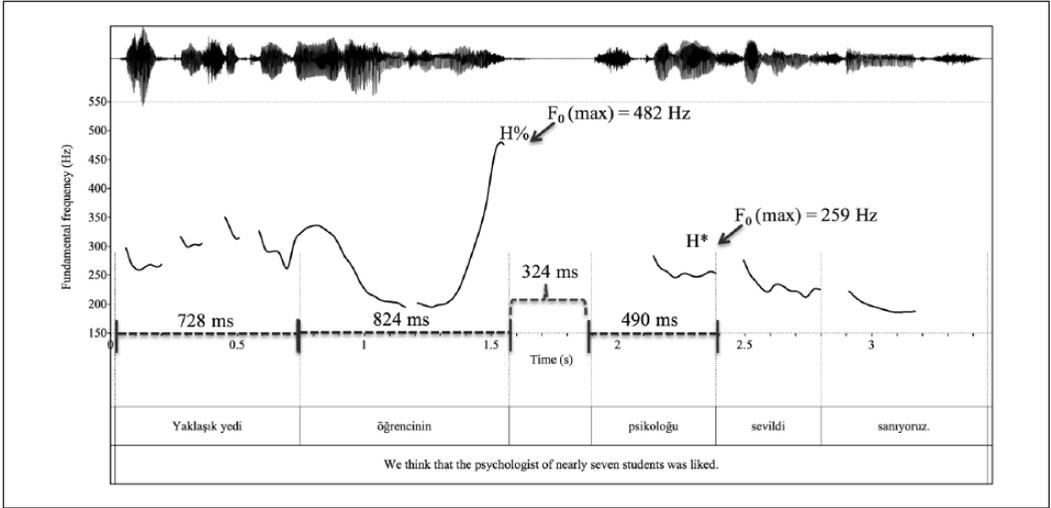
Figure 3. Waveform and pitch track for one LC sentence ((6c) above) uttered with conflicting prosody.
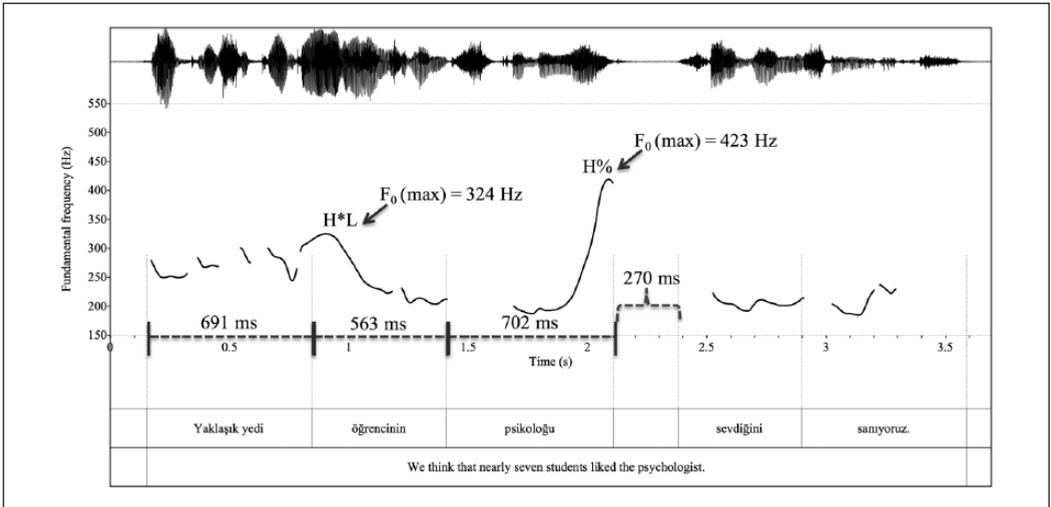

Figure 4. Waveform and pitch track for one EC sentence ((6d) above) uttered with conflicting prosody.

There were 24 experimental sentence sets, each set in twelve conditions manipulating length (lengthened subject and lengthened VP), prosody (cooperating, conflicting and neutral) and syntax (LC, EC). Length was a between subjects condition. Each length condition had six lists counterbalancing for the three prosody types and the two syntactic structures. Thus, each list included 24 target sentences. In each list, there were also 24 sentences of a different ambiguity with cooperating, conflicting and neutral prosody, and 48 unambiguous fillers of various syntactic construction types, of which 24 had neutral prosody and 24 had congruent prosody. (Dinçtopal-Deniz, 2014, provides further details on the other ambiguity and the fillers). In addition, there
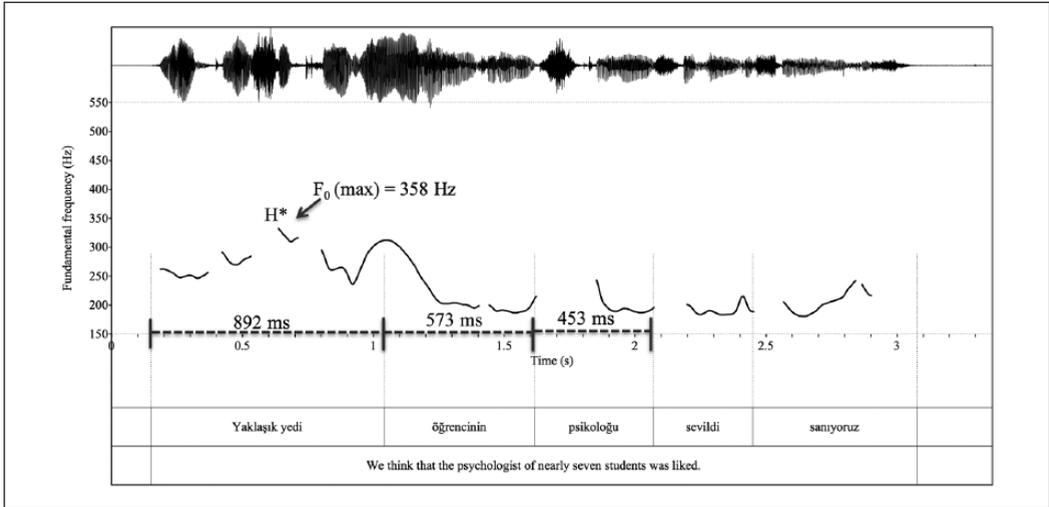
Figure 5. Waveform and pitch track for one LC sentence ((6e) above) uttered with neutral prosody.
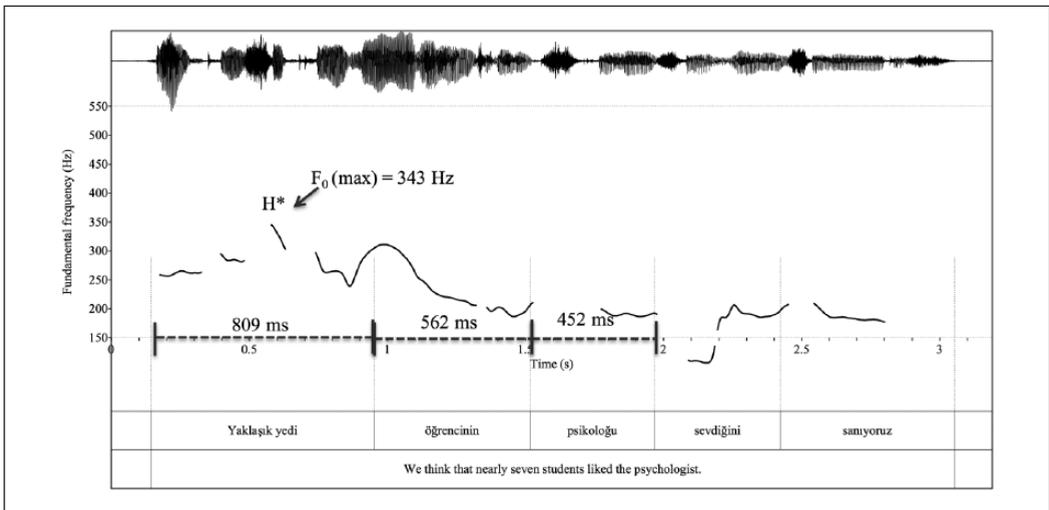
Figure 6. Waveform and pitch track for one EC sentence ((6f) above) uttered with neutral prosody.

were 10 items used in a practice session prior to the beginning of the experiment and 10 'warm-up' filler items, 5 at the beginning of each list and 5 half way through, where participants were encouraged to take a rest break. Thus, each list following the practice session had a total of 106 sentences.

    In the lengthened subject condition, the cooperating prosody conditions yielded non-optimal prosodic phrasing (4+2 PWds) for the LC structures and an optimal prosodic phrasing (3+3 PWds) for the EC structures. The conflicting prosody conditions, on the other hand, yielded an optimal prosodic phrasing (3+3 PWds) for the LC syntax but non-optimal prosodic phrasing (4+2 PWds)

Table 1. Between- and within-subjects conditions in the study.

Between-subjects conditions	Within-subjects conditions		
	Cooperating prosody	Conflicting prosody	Neutral prosody
Lengthened subject	LC (4+2 PWds)	LC (3+3 PWds)	LC (6 PWds)
	EC (3+3 PWds)	EC (4+2 PWds)	EC (6 PWds)
Lengthened VP	LC (3+3 PWds)	LC (2+4 PWds)	LC (6 PWds)
	EC (2+4 PWds)	EC (3+3 PWds)	EC (6 PWds)

for the EC structures. In the neutral prosody condition, there was no disambiguating prosodic boundary; there was just one long prosodic phrase with six PWds, which did not sound unnatural due to the limited range of prosodic variation following the focused phrase. It is important to note, however, that dividing such a sentence into two prosodic phrases is also perfectly natural in Turkish. If it were mentally divided by listeners, it would naturally tend to break in line with the syntactic phrasing. That division would be into 4+2 in the LC disambiguated condition and into 3+3 in the EC disambiguated condition.

In the lengthened VP condition, the cooperating prosody conditions yielded an optimal prosodic phrasing (3+3 PWds) for the LC syntax and non-optimal prosodic phrasing (2+4 PWds) for the EC syntax. The conflicting prosody conditions, on the other hand, yielded non-optimal prosodic phrasing (2+4 PWds) for the LC syntax and an optimal prosodic phrasing (3+3 PWds) for the EC syntax. In the neutral prosody condition there was no prosodic boundary signaling either the correct or incorrect structure but the natural syntactic division was 3+3 in the LC condition and 2+4 in the EC condition.

Table 1 summarizes the between-subjects conditions (lengthened subject, lengthened VP) and within-subjects conditions (prosody: cooperating, conflicting, neutral; and syntax: LC, EC).

2.4.4 Procedure. The experiment used a timed end-of-sentence comprehension 'got it' task. In this task, participants are asked to indicate, as quickly as possible after hearing each sentence, whether or not they have understood it, by pressing one of two keys on the keyboard. At random intervals a sentence is followed by a comprehension question to ensure attention. Reaction time (RT) for 'got it' responses is taken as a measure of the ease or difficulty of processing the sentence.

The 'got it' task taps comprehensibility judgments sentence-finally, which was important so that the participants would have full knowledge of the phrase lengths when making their judgment. This task has been used in other experiments on syntactic parsing (e.g., Frazier, Clifton, & Randall, 1983; Kjelgaard & Speer, 1999). For working with spoken language, it has the advantage of yielding an immediate response to the stimulus before any intrusion of other material such as comprehension questions (as in Clifton et al., 2006). Reading a question, even silently, may interfere with the memory of the target utterance, including its prosodic properties. This task also avoids drawing listeners' attention to the presence of ambiguity, whereas that is difficult to achieve if two competing paraphrases are presented to participants.

Following informed consent procedures, a participant was seated comfortably in front of a computer in a quiet room. The sentences were presented auditorily via noise-cancelling headphones. Participants were given instructions by the researcher at the beginning of the experiment. They were told to listen to the sentences carefully and at the end of each sentence to

indicate as quickly as possible whether or not they had comprehended the sentence, by pressing either the 'yes' button or the 'no' button (written on a green and red background respectively) on the keyboard. They were also instructed that they would be visually presented with comprehension questions following some sentences, so they needed to listen to all the sentences carefully in case a comprehension question would follow. There were 24 comprehension questions, which tapped the sentence content but did not draw attention to ambiguities in the target items. Half the questions followed experimental items and the other half followed filler items. The questions appeared on the screen, immediately after the participant's yes or no 'got it' response. The practice session items and warm-up sentences were also followed by intermittent comprehension questions.

2.4.5   Predictions. We begin with the cooperating and conflicting prosody conditions, in which there was an IPh boundary. The RSH predicts that cooperative prosodic phrasing will be most supportive for syntactic processing when the boundary flanks a short constituent (i.e., in 2+4 and 4+2) and hence cannot be attributed to purely rhythmic considerations. And it predicts that conflicting prosodic phrasing will be least disruptive for syntactic processing when it flanks more typical length constituents (as in 3+3), and hence *can* be attributed to purely metric influence. In the lengthened subject condition the LC syntax is at an advantage over EC syntax in both regards (most supportive when cooperating; least damaging when conflicting). In the lengthened VP condition, by contrast, the EC syntax has the advantage over LC syntax in both of these regards. Thus, if RSH is operative in the parsing of this Turkish construction, then – all else being equal – it would result in shorter 'got it' RTs for LC items than for EC items in both cooperating and conflicting prosody in the lengthened subject condition, and shorter 'got it' RTs for EC items than for LC items in both cooperating and conflicting prosody in the lengthened VP condition.

However, if RSH does not apply what would be expected is either no consistent bias or a traditional syntax-based preference for the LC structure. Unlike the RSH, the latter would predict shorter 'got it' RTs for LC items than for EC items in both length conditions, regardless of phrase lengths patterns.

It is also possible that both prosodic and syntactic pressures are operative. In that case their relative strengths are hard to anticipate, but there is a clear prediction that the shortest RTs should obtain wherever the two influences both cooperate with the correct syntactic structure (as for LC items in the lengthened subject condition), and the longest RTs should be observed wherever the two influences both conspire against the correct syntactic structure (as for EC items in the lengthened subject condition); where the two oppose each other (as in both LC and EC items in the lengthened VP condition), intermediate RTs may be observed.

We turn now to the neutral prosody conditions. These raise quite different considerations. In the neutral prosody conditions, there was no disambiguating prosodic boundary; there was one prosodic phrase containing six PWds. The lack of an overt prosodic boundary does not violate phrase length constraints because it is standard for the prosodic contour to be 'flattened' following a contrastively stressed constituent.

In a narrow sense, the RSH does not apply to this case. But in the spirit of the original motivation for RSH, we may consider how a listener in the habit of registering whether a speaker's prosody is rationally motivated might respond to the no-boundary pronunciation. One possibility of interest is that the listener posits an underlying prosodic boundary that has been suppressed by the reduction of prosodic range following the contrastive focus. There are findings which indicate that listeners mentally project prosodic boundaries in appropriate locations, where no overt

acoustic evidence for them is present in the input (Pauker, Itzhak, Baum, & Steinhauer, 2011). In Pauker et al.'s event-related potential data, there is even a hint that what is projected is the least 'marked' pattern, which conforms best to the length constraints of the language. We may speculate, therefore, that listeners may (though they need not) project a 3+3 phrasing pattern in the neutral prosody condition. This would result in a preference for EC in the lengthened subject condition (where syntactic phrasing is 3+3) and a preference for LC in the lengthened VP condition (also syntactically 3+3). Just as for the disambiguating prosody conditions discussed above, it is possible that in the neutral prosody condition the syntactically based LC strategy may add its bias to the listener's judgment, opposing any EC preference in the lengthened subject condition and reinforcing any LC preference in the lengthened VP condition.

2.5 Data analysis and results

One participant's data were excluded due to failure to meet the criterion of >85% accuracy on the comprehension questions. Some additional data points were excluded from the analyses due to either failure to press a key before the time-out limit (20 seconds) or too quick key presses (before the sound file ended); together, these amounted to 1.4% of the data.

The data were analyzed using the R statistical computing software, version 2.15.2 (R Core Team, 2012). The RTs were first inspected for normal distribution. The analyses showed that the data did not distribute normally ($W = 0.72$, $p < 0.001$, $D = 0.18$, $p < 0.001$) and were therefore log-transformed following Baayen and Milin (2010) and Ratcliff (1993). Any outliers that were above/below ±1.5 x interquartile range were excluded from the log-transformed data (2.5% of the data) and the data showed better normal distribution after these steps ($W = 0.99$, $p < 0.01$, $D = 0.02$, $p = 0.41$). The log-transformed RTs for positive 'got it' responses were analyzed via mixed effects modeling (Baayen, Davidson, & Bates, 2008) using the *lmer* function of the lme4 package (Bates, Maechler, Bolker, & Walker, 2015). Because mixed effects modeling does not require prior averaging, it allows researchers to examine effects that unfold during the course of an experiment (Baayen et al., 2008; Baayen, 2008). For the present study, longitudinal effects of familiarization or fatigue (i.e., RTs becoming shorter or longer in time respectively) were examined to detect any noise they might bring into the data. An analysis for the relationship between the RTs and trial showed that the RTs became gradually shorter towards the end of the experiment ($\beta = -3.84$, $SE = 0.71$, $t = -12.84$, $p < 0.001$). Thus, in building a model for the analyses, trial (after being centered to prevent collinearity (Baayen, 2008)) was included as one of the predictor variables.

The analyses were run on log-transformed RTs. Length, prosody and syntax were fit as fixed factors and subject and items as random factors. Analyses started with simpler models and were built up to where length (lengthened subject, lengthened VP), prosody (cooperating, conflicting and neutral) and syntax (LC and EC) interacted. The model with the three-way interaction of length, prosody and syntax was compared via a likelihood ratio test to a simpler model including the prosody and syntax interaction but not including length. This indicated that the model with the three-way interaction explained the data better than the simpler model with the two-way interaction, $\chi^2(6) = 13.4$, $p < 0.05$.

The three-way interaction indicates that phrase lengths do influence the perceived informativeness of prosodic boundaries but that this is affected by the type of prosody (cooperating, conflicting and neutral) and by the syntactic structure (LC, EC). The following section presents analyses for the lengthened subject and lengthened VP conditions, to better understand the nature of this interaction.
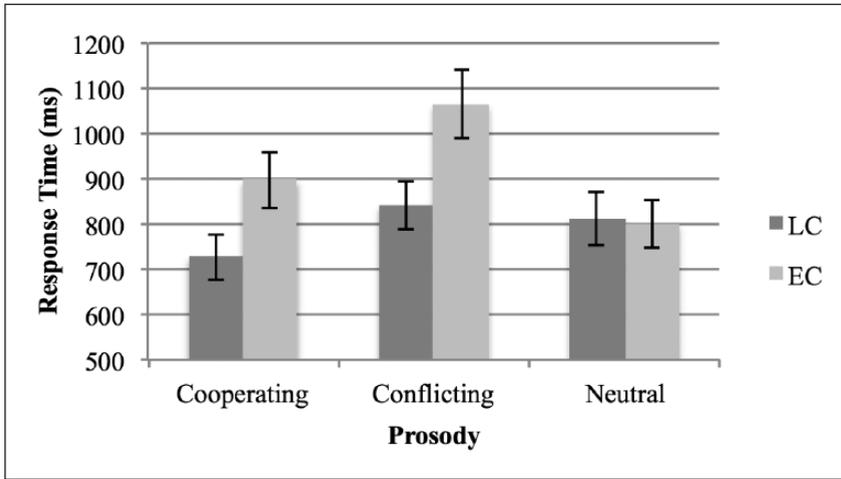
Figure 7. Lengthened subject condition: mean response times with standard errors for 'understood' responses.

2.5.1 Within-subjects analyses

2.5.1.1 Lengthened subject condition. As in the case of the between-subjects data, an analysis for the relationship between the RTs and trial in the lengthened subject condition showed that the RTs became gradually shorter towards the end of the experiment ($\beta = -3.8$, $SE = 0.44$, $t = -8.67$, $p < 0.001$). Thus, trial (after being centered to prevent collinearity) was included as one of the predictor variables and was adjusted to vary by-subject. The model with by-subject adjustment explained the data better than the one without any by-subject adjustments, $\chi^2(2) = 17.56$, $p < 0.001$.

Prosody and syntax were fit as fixed factors and subjects and items as random factors. As for the combined data, analyses started with simpler models and were built up to where prosody and syntax interacted. A likelihood-ratio test comparing the complex model with interaction to the simpler models showed that the model including the interaction explained the data better than the simpler ones, $\chi^2(3) = 8.03$, $p < 0.05$. This interaction model also allowed for random slopes for subjects by prosody, $\chi^2(7) = 19.56$, $p < 0.01$.

There were also main effects of prosody and syntax. The model for prosody showed that sentences with conflicting prosody were processed slower than those with neutral prosody ($\beta = 102$, $SE = 31.35$, $t = 3.59$, $p < 0.001$). But there was no reliable difference between sentences with cooperating prosody and those with neutral prosody ($\beta = -15.31$, $SE = 25.71$, $t = -0.6$, $p = 0.95$). The model with syntax as a predictor variable showed that the LC structures were processed faster than the EC structures ($\beta = -56.32$, $SE = 21.82$, $t = -2.51$, $p < 0.05$).

The model with interaction was investigated via quantile–quantile (q–q) plots, and data points with standardized residuals below/above 2.5 standard deviations were excluded from the analyses (Baayen & Milin, 2010). Following Baayen and Milin (2010), subjects, items and individual data points were inspected to identify any overly influential subjects, items, or data points, by using the *influence.ME* function (Nieuwenhuis, te Grotenhuis, & Pelzer, 2012). Both Cook's distance values and plots were used in this inspection. Two subjects, 3 items and 5 individual data points diverged from group statistics. Divergent data were excluded and the model was re-fit.

RTs from the remaining participants and items are shown in Figure 7.

Pairwise comparisons using the *glht* function showed that in cooperating and conflicting prosody conditions, structures with LC syntax were processed faster than those with EC syntax
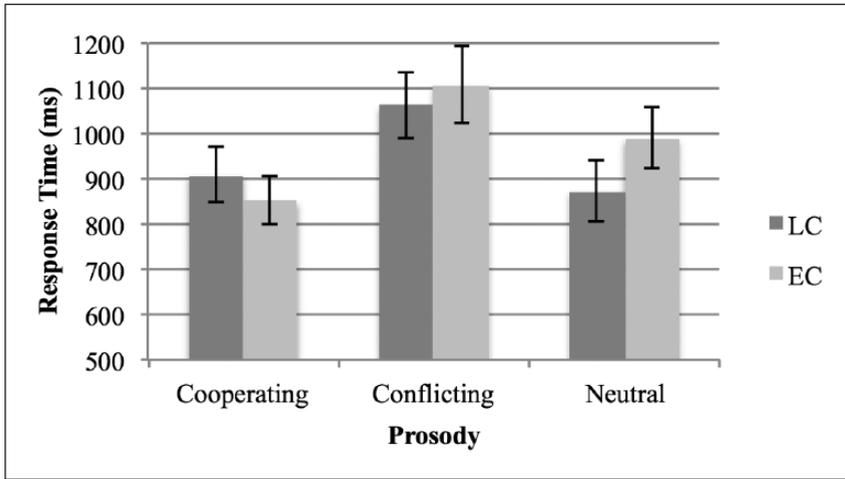
Figure 8. Lengthened VP condition: mean response times with standard errors for 'understood' responses.

($\beta = -0.125$, $SE = 0.059$, $z = -2.10$, $p < 0.05$ for cooperating prosody, and $\beta = 0.251$, $SE = 0.062$, $z = -4.03$, $p < 0.001$ for conflicting prosody). There was no significant difference between LC and EC syntax in the neutral prosody condition ($\beta = 0.00007$, $SE = 0.062$, $z = 0.001$, $p = 0.992$).

2.5.1.2 Lengthened VP condition. Analyses parallel to those for the lengthened subject condition were conducted for the lengthened VP condition. Longitudinal effects of familiarization or fatigue were examined via a mixed effects model for the RTs with trial number as the only fixed effects term. This analysis showed that the RTs became shorter towards the end of the experiment ($\beta = -4.86$, $SE = 0.48$, $t = -10.8$, $p < 0.001$). Thus, in the models for the main analyses, trial (after being centered to prevent collinearity) was included as one of the predictor variables and was adjusted to vary by subject. The model allowing for random slopes for trial for subjects was significantly better than those with random intercepts only, $\chi^2(2) = 26.36$, $p < 0.001$.

The main analyses were run on the logRT with prosody and syntax as the independent variables. While building a model, each predictor variable was first entered into the model separately; a more complex model including the two predictors was then built. A likelihood ratio analysis comparing the simple models to the complex one with interaction indicated that the model including both predictors accounted for the data better than the simpler ones ($\chi^2(1) = 0.51$, $p = 0.47$ for the comparison of the simple model for prosody vs. the complex model with prosody and syntax; $\chi^2(2) = 44.64$, $p < 0.001$ for the comparison of the simple model for syntax vs. the complex model with prosody and syntax). The complex model also included random slopes for subjects by prosody, $\chi^2(7) = 16.03$, $p < 0.05$.

There was also a main effect for prosody which indicated that sentences with cooperating prosody were processed faster than those with neutral prosody ($\beta = -0.52.41$, $SE = 25.87$, $t = -2$, $p < 0.05$) and sentences with conflicting prosody were processed slower than those with neutral prosody ($\beta = 138.96$, $SE = 33.88$, $t = 4.58$, $p < 0.001$). The model with syntax did not indicate any significant difference between the LC and EC structures overall ($\beta = -17.96$, $SE = 23.84$, $t = -0.76$, $p = 0.45$).

During model criticism for the complex model with prosody and syntax interaction, data points with standardized residuals below/above 2.5 standard deviations were excluded from the analyses, as were 4 overly influential subjects, 4 items and 3 individual data points. The remaining RTs are shown in Figure 8.

Table 2. Percent 'understood' responses.

Between Subjects Conditions		Cooperating Prosody (98)	Conflicting Prosody (92)	Neutral Prosody (93)
Lengthened Subject	**LC**	96	97	94
	EC	98	85	92
Lengthened VP	**LC**	99	94	96
	EC	98	92	90

Planned pairwise comparisons showed that in cooperating and conflicting prosody conditions, there was no reliable difference between the LC syntax and the EC syntax structures ($\beta = 6.97$, $SE = 44.77$, $z = -0.161$, $p = 0.87$ for cooperating prosody, and $\beta = -2.19$, $SE = 45.08$, $z = -0.05$, $p = 0.96$ for conflicting prosody). But for the neutral prosody condition, the LC structures were processed faster than the EC structures ($\beta = -86.12$, $SE = 39.63$, $z = 2.1$, $p < 0.05$).

Discussion of these findings will follow a brief presentation of the 'understood' responses (i.e., 'yes' button press to the 'got it?' query). Table 2 presents the percentage of overall 'understood' responses and the percentage of 'understood' responses in each between-subjects condition.

As can be seen from Table 2, the percentages of 'understood' responses were quite high (above 85%) for all the conditions. This is as expected. The main purpose of the 'got it' task is to assess how quickly comprehension is achieved.

The mixed effects logistic regression analyses for the lengthened subject condition indicated that participants' tendency to respond 'yes' was higher for the cooperating prosody when compared with the neutral prosody (odds ratio: $\beta = 2.6$, $SE = 1.47$, $z = 2.47$, $p < 0.05$). There was no significant difference in 'yes' responses between the conflicting and neutral prosody conditions ($\beta = 0.67$, $SE = 1.35$, $z = -1.3$, $p = 0.18$). Participants' responses to the two syntactic constructions (LC/EC) in the lengthened subject condition show a pattern parallel to their processing times (see Figure 7) for both conflicting and neutral prosody. That is, in the conflicting prosody condition, participants were more likely to respond 'yes' for LC syntax than EC syntax ($\beta = 11.1$, $SE = 1.7$, $z = 4.51$, $p < 0.001$), while in the neutral prosody condition there was no significant difference in the probability of 'yes' responses for LC and EC syntax ($z = 0.45$). However, in the cooperating prosody condition, the 'yes' responses (unlike the RT data) did not differ significantly between LC and EC syntax ($z = 1.61$), possibly due to ceiling effects.

For the lengthened VP condition, analyses showed that participants were more likely to respond positively to the cooperating prosody than to the neutral prosody (odds ratio: $\beta = 6.24$, $SE = 1.67$, $z = 3.54$, $p < 0.001$), but the conflicting prosody showed no reliable difference compared with the neutral prosody ($\beta = 1.04$, $SE = 1.38$, $z = 0.13$, $p = 0.89$). Mirroring the RT data, there was no reliable difference in positive responses between LC and EC syntax in the cooperating and conflicting prosody conditions (z's < 1.4) but participants were more likely to indicate that they had understood the LC syntax than the EC syntax in the neutral prosody condition ($\beta = 4.3$, $SE = 1.68$, $z = 2.82$, $p < 0.05$).

2.6 Discussion

Our discussion of the results will focus exclusively on the RT data since the percent 'understood' data were limited by ceiling effects, as is not uncommon for 'got it' methodology. Table 3 presents a summary of the RT data findings in both the lengthened subject and lengthened VP conditions.

As Table 3 illustrates, in the lengthened subject condition there was an LC advantage in cooperating and conflicting prosody conditions. In the neutral prosody condition, there was no

Table 3. Lengthened subject and lengthened VP conditions, summary of response time data pattern. '<' indicates faster processing, '=' indicates no significant difference in processing time. All inequalities in the table are confirmed at $p < 0.05$ or smaller. For phrase lengths in each of these conditions, see Table 1.

	Cooperating prosody	Conflicting prosody	Neutral prosody
Lengthened subject	LC < EC	LC < EC	LC = EC
Lengthened VP	LC = EC	LC = EC	LC < EC

advantage for either structure. In the lengthened VP condition, there was no LC advantage in the cooperating and conflicting prosody conditions, though there was an LC advantage in the neutral prosody condition.

Considering just the cooperating and conflicting conditions first, the important finding is that the LC advantage in the lengthened subject condition was absent in the lengthened VP condition. This implies that the LC advantage observed in the lengthened subject condition was not solely due to a syntactic Late Closure strategy. Phrase lengths evidently also influenced the ease of processing, presumably indirectly through their effect on how prosodic boundaries are evaluated by listeners as indicators of syntactic structure, as proposed by the RSH. However, phrase lengths alone are insufficient to explain the total data set, since the phrase lengths in the lengthened VP condition would have favored EC, yet the results showed no EC advantage in the cooperating and conflicting prosody conditions. To account for the absence of an advantage for either LC or EC in these conditions, it seems necessary to suppose: (i) that there was also some general pressure towards the LC structure, such as would stem from a syntactic Late Closure strategy; and (ii) that there was a trade-off in the lengthened VP condition between the RSH, favoring EC, and syntactic Late Closure, which would disfavor EC. This differs importantly from the findings for the lengthened subject condition, where the RSH would favor LC, and there would be no conflict with a syntactic Late Closure bias.

Continuing to explore this possibility that there is interplay between syntactic Late Closure and phrase lengths, we consider now the neutral prosody condition, which showed an advantage for the LC structure in the lengthened VP condition but no reliable difference between LC and EC in the lengthened subject condition. The RSH is inapplicable in this condition with no overt prosodic boundary. The observed preference for the LC structure in the lengthened VP condition could be due to the syntactic Late Closure strategy, and/or to a tendency to mentally project a prosodic boundary (Pauker et al., 2011), most probably where it would create an optimal pattern of prosodic phrase lengths (3+3). In fact both of these factors are essential for explaining the total pattern of data, particularly the case of neutral prosody in the lengthened subject condition. There, the Late Closure strategy would favor LC structure while projection of an optimal prosodic phrasing would favor EC structure. The fact that there was no structural advantage at all in that case suggests that these two factors are more or less equal in strength and cancelled each other out.

3   General discussion

This study investigated three main questions, not previously addressed in the literature on the prosody–syntax interface: (i) Are speakers of Turkish, a language whose structure is very different from English, influenced by phrase lengths in their assessment of the informativeness of prosodic cues, as shown for English by Clifton et al. (2006)? (ii) If so, do such length considerations play a role in listeners' evaluation of prosodic cues that conflict with the syntax, as well as syntactically appropriate prosodic cues? (iii) How do listeners weigh interactions between syntactically defined parsing strategies and metrically based factors in their evaluation of prosodic cues to syntactic structure?

The results of the study provided a positive answer for the first two questions: Turkish speakers are influenced by phrase lengths in their interpretation of prosodic cues in line with the RSH, and this is so not only when prosodic breaks are aligned with syntactic boundaries but also when they are not. The answer to the third question is that both factors have roughly equal weight. The results showed that for Turkish speakers, at least for the ambiguity investigated here, a structure favored by the Late Closure strategy but disfavored by phrase lengths as per the RSH is not significantly easier or harder to process than a structure disfavored by the Late Closure strategy but favored by phrase lengths as per the RSH. When they are in competition, the two biases essentially cancel each other out.

In terms of the general conclusions that can be drawn, the finding that the Late Closure strategy is alive and well is not especially remarkable, since the construction we tested involves primary relations, and Late Closure (or close relatives of it such as low attachment, local attachment, and recency) is widely accepted as operative in the parsing of primary relations, even by those such as Frazier and Clifton (1996) who reject Late Closure for non-primary relations. For the RSH, on the other hand, some novel clarification has been gained from this study. Ours is the first demonstration of RSH at work in the parsing of primary relations. Clifton et al. (2006) tested RSH only for non-primary relations in ambiguities of adverb phrase attachment (or 'association' in Construal Theory) and scope of coordination.

The nature of the interplay between Late Closure and RSH also provides new insight. Though their observable effects do indeed overlap in some contexts, each of these parsing strategies has its own territory: Late Closure is sensitive to height in syntactic tree structure, and RSH is sensitive to the sizes of sequential prosodic phrases, and this is why they can be distinguished. Of particular interest is that where their effects oppose each other, the two strategies were found (at least for our materials) to be on a more or less equal footing in terms of their relative strength of influence on the parser's decision. This implies that RSH is a quite powerful factor, on a par with the generally accepted strong influence of Late Closure. Together, these two observations (application to primary relations, and strong impact) suggest that RSH may be an integral part of the sentence processing machinery and should be considered as a factor in future parsing studies, with phrase lengths controlled or manipulated with care.

Our results also have possible implications for the characterization of the prosodic phenomena which drive the RSH. So far we have characterized the relevant difference in these Turkish materials, a difference between 3+3 prosodic phrasing and 4+2 or 2+4, as due to the 'marked' quality in Turkish of short prosodic phrases of less than 3 prosodic words. However, it could be viewed instead in terms of balanced vs. unbalanced (uniform vs. non-uniform) prosodic phrasing. A general expectation of length-balanced phrasing would explain the same findings as above: the 4+2 and 2+4 patterns being taken more seriously than 3+3 in the cooperating condition and being more difficult to ignore in the conflicting condition. The balance explanation would also apply to the neutral prosody (no boundary) items, which favored the 3+3 phrasing. Our experiment was not designed to distinguish these alternative descriptions of the prosodic contours. For that it will be necessary to use sentences with more varied prosodic phrases. For example, are patterns such as 3+4 (unbalanced but within natural length limits in Turkish) and 2+2 (balanced but non-optimal lengths) as natural as 3+3 in Turkish?

Characterization in terms of balance makes contact with a broader literature on prosodic influences on syntactic parsing. A tendency towards balanced or uniform prosodic phrasing has been reported for several languages in the linguistic literature. It was foreshadowed by Gee and Grosjean (1983) for English and given further attention in work by Ghini (1993) for Italian, subsequently followed up by Sandalo and Truckenbrodt (2002) with data from Brazilian Portuguese. Pynte (2006) provides extensive psycholinguistic evidence for uniformity in listeners' attachment of prepositional phrases in French. He investigated a temporary ambiguity involving a verb modifier

or a noun modifier interpretation of a prepositional phrase (PP). He observed that French listeners relied on the lengths of previous prosodic units to predict the length of a following one, generating syntactic expectations. When a long prosodic phrase (containing a post-verbal NP PP sequence) followed a series of short prosodic phrases, listeners regarded its greater length as motivated by the syntactic structure: it was most often parsed as a single syntactic constituent with the PP modifying the NP (*removed [the chain of the bicycle]*). By contrast, when the NP PP prosodic phrase followed prosodic phrases that were just as long, it carried no such syntactic implications: the PP was often understood to be a separate element within the VP (*removed [the chain] [from the bicycle]*). Thus, Pynte's data suggest that listeners take the uniformity of constituents in French into account in assigning syntactic structure. When the absence of a break in a long phrase is not explicable in terms of prosodic uniformity it is taken seriously as a cue to syntactic disambiguation.

Prosodic uniformity is not specifically mentioned in Clifton et al.'s (2006) discussion of RSH, which distinguishes 'short' versus 'long' prosodic phrases (in terms of their number of syllables). This is reflected in one formulation of RSH: "prosodic boundaries will have a larger influence on listeners' choice of an analysis when they flank short constituents than when they flank long ones" (Clifton et al., p. 854). But within the general concept of the RSH, it is imaginable that other metric constraints, including uniformity, may also be taken into account by listeners in assessing the significance of a prosodic boundary. This would still be very much in the general spirit of RSH, as expressed in Clifton et al.'s broader formulation: "listeners can be sensitive to the demands placed on speakers and take these demands into account in determining speakers' intentions" (Clifton et al., p.858). Indeed, Clifton et al. (2006) themselves draw in another prosodic phenomenon under the RSH umbrella. They refer to their earlier finding that "the interpretation of a prosodic boundary is determined not by its absolute size but by its size relative to relevant certain other boundaries." This is the Informative Boundary Hypothesis of Clifton et al. (2002, p. 87). It is imaginable that future investigations of prosody–syntax interface phenomena might ultimately lead to the conclusion that *any* potentially syntactically relevant prosodic property of an utterance is likely to be evaluated by listeners as reflecting the inferred intentions of the speaker.

Acknowledgements

Funding

References

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge, UK: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, *3*(2), 12–28.

Bates, D., Maechler, M., Bolker, B. M., & Walker, S. (2015). Fitting Linear Mixed-Effects Models using {lme4}. *Journal of Statistical Software*, *67*(1), 1–48. http://doi.org/10.18637/jss.v067.i01

Carlson, K., Clifton, C., & Frazier, L. (2001). Prosodic boundaries in adjunct attachment. *Journal of Memory and Language*, *45*(1), 58–81. http://doi.org/10.1006/jmla.2000.2762

Carlson, K., Clifton, C., & Frazier, L. (2009). Nonlocal effects of prosodic boundaries. *Memory & Cognition*, *37*(7), 1014–1025. http://doi.org/10.3758/mc.37.7.1014

Clifton, C., Carlson, K., & Frazier, L. (2002). Informative prosodic boundaries. *Language and Speech*, *45*(2), 87–114. http://doi.org/10.1177/00238309020450020101

Clifton, C., Carlson, K., & Frazier, L. (2006). Tracking the what and why of speakers' choices: Prosodic boundaries and the length of constituents. *Psychonomic Bulletin & Review*, *13*(5), 854–861.

Dinçtopal-Deniz, N. (2014). *The interplay of syntactic processing strategies and phrase lengths in processing Turkish sentences*. PhD Dissertation. Graduate Center, City University of New York, New York. Retrieved from http://academicworks.cuny.edu/gc_etds/346/

Erguvanlı, E. E. (1984). *The function of word order in Turkish grammar*. Berkeley, CA: University of California Press.

Frazier, L. (1978). *On comprehending sentences: Syntactic parsing strategies*. PhD Dissertation. University of Connecticut, Storrs. retrieved from https://www.researchgate.net/publication/27401728_On_Comprehending_Sentences_Syntactic_Parsing_Strategies

Frazier, L., Carlson, K., & Clifton, J. C. (2006). Prosodic phrasing is central to language comprehension. *Trends in Cognitive Sciences*, *10*(6), 244–249. http://doi.org/10.1016/j.tics.2006.04.002

Frazier, L., & Clifton, C. (1996). *Construal*. Cambridge, MA: MIT Press.

Frazier, L., Clifton, C., & Randall, J. (1983). Filling gaps: Decision principles and structure in sentence comprehension. *Cognition*, *13*(2), 187–222. http://doi.org/10.1016/0010–0277(83)90022–7

Gee, J. P., & Grosjean, F. (1983). Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology*, *15*(4), 411–458. http://doi.org/10.1016/0010–0285(83)90014–2

Ghini, M. (1993). Ø-formation in Italian: A new proposal. In C. Dyck (Ed.), *Toronto Working Papers in Linguistics* (Vol. 12 (2), pp. 41–78). Toronto, Canada: Department of Linguistics, University of Toronto.

Hwang, H., & Schafer, A. J. (2009). Constituent length affects prosody and processing for a dative NP ambiguity in Korean. *Journal of Psycholinguistic Research*, *38*(2), 151–175. http://doi.org/10.1007/s10936–008–9091–1

Inkelas, S., & Orgun, C. O. (2003). Turkish stress: A review. *Phonology*, *20*(1), 139–161.

Ipek, C. (2011). Phonetic realization of focus with no on-focus pitch range expansion in Turkish. *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS 2011)*, 140–143. Retrieved from http://www.homepages.ucl.ac.uk/~uclyyix/focus_crosslinguistic/Ipek_ICPhS2011.pdf

Kamali, B. (2008). Scope of negation and phonological phrasing in Turkish. *Paper presented at Mediterranean Syntax Meeting II (MSM-II)*. Boğaziçi University.

Kjelgaard, M. M., & Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language*, *40*(2), 153–194. http://doi.org/10.1006/jmla.1998.2620

Lehiste, I. (1973). Phonetic disambiguation of syntactic ambiguity. *Glossa*, *7*(2), 107–122.

Marslen-Wilson, W. D., Tyler, L. K., Warren, P., Grenier, P., & Lee, C. S. (1992). Prosodic effects in minimal attachment. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *45*(1), 73–87.

Nagel, H. N., Shapiro, L. P., Tuller, B., & Nawy, R. (1996). Prosodic influences on the resolution of temporary ambiguity during on-line sentence processing. *Journal of Psycholinguistic Research*, *25*(2), 319–344. http://doi.org/10.1007/bf01708576

Nash, R. (1973). *Turkish intonation*. The Hague, The Netherlands: Mouton de Gruyter.

Nieuwenhuis, R., te Grotenhuis, M., & Pelzer, B. (2012). influence.ME: Tools for detecting influential data in mixed effects models. *The R Journal*, *4*(2), 38–47.

Pauker, E., Itzhak, I., Baum, S. R., & Steinhauer, K. (2011). Effects of cooperating and conflicting prosody in spoken English garden path sentences: ERP evidence for the boundary deletion hypothesis. *Journal of Cognitive Neuroscience*, *23*(10), 2731–2751. http://doi.org/10.1162/jocn.2011.21610

Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America*, *90*(6), 2956–2970.

Pynte, J. (2006). Phrasing effects in comprehending PP constructions. *Journal of Psycholinguistic Research*, *35*(3), 245–265. http://doi.org/10.1007/s10936–006–9014–y

R Core Team. (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3–900051–07–0. Retrieved from http://www.R-project.org/

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*(3), 510–532.

Sak, H., Güngör, T., & Saraçlar, M. (2008). Turkish language resources: Morphological parser, morphological disambiguator and web corpus. *GoTAL*, *5221*, 417–427.

Sandalo, F., & Truckenbrodt, H. (2002). Some notes on phonological phrasing in Brazilian Portuguese. *MIT Working Papers in Linguistics*, *42*(1), 285–310.

Schafer, A. J., Speer, S. R., Warren, P., & White, S. D. (2000). Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research*, *29*(2), 169–182.

Speer, S. R., Kjelgaard, M. M., & Dobroth, K. M. (1996). The influence of prosodic structure on the resolution of temporary syntactic closure ambiguities. *Journal of Psycholinguistic Research*, *25*(2), *249–271*. http://doi.org/10.1007/bf01708573

Stoyneshka, I., Fodor, J. D., & Fernández, E. M. (2010). Phoneme restoration methods for investigating prosodic influences on syntactic processing. *Language and Cognitive Processes*, *25*(7), 1265–1293.

Vallduví, E., & Engdahl, E. (1996). The linguistic realization of information packaging. *Linguistics*, *34*(3), 459–519.

Webman Shafran, R. (2011). *Prosody and parsing in a double PP construction in Hebrew*. PhD Dissertation. City University of New York, New York.

Appendices

Appendix A. Pre-tests

Appendix A.1: Corpus data and normative study. Noun phrase (NP) sequences which had a potentially ambiguous morphological ending on the second NP (3SG.POSS or ACC; i.e., LC or EC interpretation) were elicited from a written Turkish corpus of 17 million words (Sak, Güngör, & Saraçlar, 2008). Forty NP sequences, matched for the log frequency of the alternative morphological endings of the second NP (3SG.POSS or ACC), were selected to be used in creating a pool of potential experimental sentences. Forty sentences (with LC and EC continuations after the temporarily ambiguous NP), each in two length versions (lengthened subject vs. lengthened VP) were created using these corpus-based ambiguous NP sequences, by the addition of appropriate adjective phrases (e.g., '(yaklaşık) yedi') and verb phrases (e.g., '(oldukça) sevildi/sevdiğini'); see examples (6) and (7) above.

A normative study following the method in Webman Shafran (2011) was used to pretest these potential experimental sentences to eliminate ones that were strongly biased towards one interpretation. The length versions were distributed across two reading lists to be judged on a 7-point Likert scale by four judges who were native speakers of Turkish with a background in linguistics (please see further details in Dinçtopal-Deniz, 2014).

The 24 items which were the least biased towards an LC or an EC interpretation were selected for use in the main experiments. The lengthened subject condition received a mean score of 3.83 and the lengthened VP condition received a mean score of 3.9; the difference between the two conditions was not reliable: $t(23) = 0.923$, $p = 0.366$. Thus, both syntactic analyses of the potentially ambiguous NP sequence were semantically acceptable, in both length versions, in the sentences selected for the main experiment. See Appendix B for a list of the items used in the main experiments.

Appendix A.2. Pronunciation acceptability. The 24 items selected on the basis of the normative study were read aloud and recorded by the first author (a native speaker of Turkish). Twelve expert judges (linguistically informed native speakers of Turkish) rated the acceptability of the items in the three prosody conditions (cooperating, conflicting, and neutral). Cooperating prosody sentences in both length conditions were all rated as fully acceptable (100%). Neutral prosody sentences in each length condition were rated as highly acceptable (mean 98% for lengthened subject condition and 100% for lengthened VP condition), ensuring that the strategy of using contrastive focus to mute the prosody in the remainder of the sentence did not make the sentences sound unnatural. As predicted,

conflicting prosody conditions received very low acceptability ratings: mean 4% for lengthened subject and 2% for lengthened VP.

The spoken sentences were additionally analyzed acoustically to ensure that the perceived prosodic properties were consistently realized in the signal.

Appendix A.3. Phonetic analyses. Durational measures and fundamental frequency (F0) analyses were conducted on the temporarily ambiguous region to confirm that the items were pronounced with the intended prosody. It was desired that the phonetic measures would indicate significant differences between the LC/EC prosody in the cooperating and conflicting prosody conditions but that there would be no difference in the LC/EC neutral prosody conditions. Statistical comparisons (*t*-tests) confirmed this to be so. Table A.3.1 presents the phonetic measures of the temporarily ambiguous region in all three prosody conditions; the *t*-test comparisons on the temporarily ambiguous regions are presented in Tables A.3.2–A.3.4.

In addition to analyzing the prosodic properties of the syntactically ambiguous region, other regions were also inspected for possible intrusion of an additional prosodic boundary. That would be undesirable, especially as it has been observed that the impact of prosodic boundaries is not uniform and can show variation depending on the presence, size and location of certain other prosodic boundaries in a sentence (Carlson, Clifton, & Frazier, 2001, 2009; Frazier, Carlson, & Clifton, 2006). These analyses offered confidence that no boundary in addition to the disambiguating LC and EC prosodic breaks was present (please see Dinçtopal-Deniz, 2014 for details).

As an anonymous reviewer noted, EC items had longer pause durations than LC items in the cooperating prosody condition ($p < 0.001$); see Table A.3.2. This is a natural prosodic difference in Turkish. In Turkish, subjects that are not adjacent to the VP are followed by a comma in writing (in line with the topicalization facts for spoken Turkish). This comma rule is obligatory specifically when another NP follows the subject, as it does in the EC experimental sentences. In our LC sentences, by contrast, the subject is not followed by another NP (but only by an adverb), and so there would be less need for a strong post-subject boundary to avoid misparsing. In any case, the data suggest that the greater length of the pause in the EC condition did not have a strong impact on ease of processing, since an LC advantage was observed in the lengthened subject condition and no advantage for either prosody was observed in the lengthened VP condition.

Table A.3.1. Phonetic measures of the temporarily ambiguous region: average F0 maximum (Hz), average word and pause duration (milliseconds), NA = not applicable.

	Lengthened subject			Lengthened verb phrase		
	Duration N1/N2	F0 max N1/N2	Pause N1/N2	Duration N1/N2	F0 max N1/N2	Pause N1/N2
Cooperating						
LC	504/612	328/436	NA/266	508/618	341/432	NA/234
EC	650/390	449/275	340/NA	668/391	446/272	344/NA
Conflicting						
LC	666/392	451/276	376/NA	694/392	455/273	348/NA
EC	487/618	340/440	NA/369	506/635	362/433	NA/346
Neutral						
LC	458/347	287/223	NA/NA	469/348	314/242	NA/NA
EC	459/349	289/223	NA/NA	474/352	308/231	NA/NA

Table A.3.2. Cooperating prosody: summary of boundary locations and comparison of word duration (milliseconds (ms)) and average maximum F0 value (Hz) for N1 and N2 in late closure (LC) and early closure (EC) syntax conditions.

Length condition	Syntax	Boundary location	Word duration	F0 maximum	Pause duration
Lengthened subject	LC	After NP2	LC-N2 > EC-N2 $t(23) = 24.85, p < 0.001$	LC-N2 > EC-N2 $t(23) = 11.04, p < 0.001$	266 ms.
	EC	After NP1	EC-N1 > LC-N1 $t(23) = 19.48, p < 0.001$	EC-N1 > LC-N1 $t(23) = 15.87, p < 0.001$	340 ms
Lengthened VP	LC	After NP2	LC-N2 > EC-N2 $t(23) = 24.16, p < 0.001$,	LC-N2 > EC-N2 $t(23) = 18.20, p < 0.001$	234 ms
	EC	After NP1	EC-N1 > LC-N1 $t(23) = 21.30, p < 0.001$	EC-N1 > LC-N1 $t(23) = 25.27, p < 0.001$	344 ms

Table A.3.3. Conflicting prosody: summary of boundary locations and comparison of word duration (milliseconds (ms)) and average maximum F0 value (Hz) for N1 and N2 in late closure (LC) and early closure (EC) syntax conditions.

Length condition	Syntax	Boundary location	Word duration	F0 maximum	Pause duration
Lengthened subject	LC	After NP1	LC-N1 > EC-N1 $t(23) = 22.27, p < 0.001$	LC-N1 > EC-N1 $t(23) = 8.34, p < 0.001$	376 ms.
	EC	After NP2	EC-N2 > LC-N2 $t(23) = 24.97, p < 0.001$	EC-N2 > LC-N2 $t(23) = 10.71, p < 0.001$	369 ms
Lengthened VP	LC	After NP1	LC-N1 > EC-N1 $t(23) = 17.95, p < 0.001$	LC-N1 > EC-N1 $t(23) = 9.928, p < 0.001$	348 ms
	EC	After NP2	EC-N2 > LC-N2 $t(23) = 38.97, p < 0.001$	EC-N2 > LC-N2 $t(23) = 25.27, p < 0.001$	346 ms

Table A.3.4. Neutral prosody: comparison of word duration (milliseconds) and average maximum F0 value (Hz) for N1 and N2 in late closure (LC) and early closure (EC) syntax conditions.

Length condition	Syntax	Word duration	F0 maximum
Lengthened subject	LC	LC-N2 = EC-N2 $t(23) = 0.32, p = 0.76$	LC-N2 = EC-N2 $t(23) = 0.03, p < 0.98$
	EC	LC-N2 = EC-N2 $t(23) = 0.07, p = 0.94$	EC-N1 = LC-N1 $t(23) = 0.09, p = 0.93$
Lengthened VP	LC	LC-N2 = EC-N2 $t(23) = 1.39, p = 0.2$	LC-N2 = EC-N2 $t(23) = 0.72, p = 0.48$
	EC	EC-N1 = LC-N1 $t(23) = 1.07, p = 0.3$	EC-N1 = LC-N1 $t(23) = 1.12, p = 0.27$

Appendix B. Experimental sentences with English translations

		Experimental sentences	English translations
1	**LC**	(Yaklaşık) Yedi hastanın hekimi (sabah) kovuldu zannettik.	We considered that the doctor of (nearly) seven patients was fired/sent away (in the morning).
	EC	(Yaklaşık) Yedi hastanın hekimi (sabah) kovduğunu zannettik.	We considered that (nearly) seven patients fired/sent away the doctor (in the morning).
2	**LC**	(Sağdan) İkinci işçinin çocuğu (çok) sevildi sanmışlar.	They thought that the child of the second worker (from the left) was liked (much).
	EC	(Sağdan) İkinci işçinin çocuğu (çok) sevdiğini sanmışlar.	They thought that the second worker (from the left) liked the child (much).
3	**LC**	(En) Öndeki kadının bebeği (biraz) hırpalandı zannettim.	I considered that the baby of the woman in the (very) front was ill-treated (a little).
	EC	(En) Öndeki kadının bebeği (biraz) hırpaladığını zannettim.	I considered that the woman in the (very) front ill-treated the baby (a little).
4	**LC**	(Sondan) Birinci doktorun ortağı (biraz) eleştirildi zannettik.	We considered that the partner of the first doctor (from the last) was criticized (a little).
	EC	(Sondan) Birinci doktorun ortağı (biraz) eleştirdiğini zannettik.	We considered that the first doctor (from the last) criticized the partner (a little).
5	**LC**	(Sağdan) Üçüncü savcının sanığı (sabah) sorgulandı sanıyorum.	I think that the suspect of the third prosecutor (from the right) was questioned (in the morning).
	EC	(Sağdan) Üçüncü savcının sanığı (sabah) sorguladığını sanıyorum.	I think that the third prosecutor (from the right) questioned the suspect (in the morning).
6	**LC**	(Baştan) Üçüncü müdürün personeli (sabah) azarlandı zannettim.	I considered that the personnel of the third manager (from the beginning) were scolded (in the morning).
	EC	(Baştan) Üçüncü müdürün personeli (sabah) azarladığını zannettim.	I considered that the third manager (from the beginning) scolded the personnel (in the morning).
7	**LC**	(Yaklaşık) Yedi öğrencinin psikoloğu (çok) sevildi sanıyoruz.	We think that the psychologist of (nearly) seven students was (much) liked.
	EC	(Yaklaşık) Yedi öğrencinin psikoloğu (çok) sevdiğini sanıyoruz.	We think that (nearly) seven students liked the psychologist (much).
8	**LC**	(Sondan) Beşinci padişahın haremi (artık) kapatıldı sanıyorum.	I think that the harem of the fifth sultan (from the last) was closed down (already).
	EC	(Sondan) Beşinci padişahın haremi (artık) kapattığını sanıyorum.	I think that the fifth sultan (from the last) closed down the harem (already).
9	**LC**	(Sağdan) Birinci ustanın çırağı (dün) kovuldu sanıyorlar.	They think that the apprentice of the first master (from the right) was fired (yesterday).
	EC	(Sağdan) Birinci ustanın çırağı (dün) kovduğunu sanıyorlar.	They think that the first master (from the right) fired the apprentice (yesterday).
10	**LC**	(Soldan) Dördüncü yazarın editörü (yine) üzüldü sandılar.	They thought that the editor of the fourth writer (from the left) was upset (again).
	EC	(Soldan) Dördüncü yazarın editörü (yine) üzdüğünü sandılar.	They thought that the fourth writer (from the left) upset the editor (again).
11	**LC**	(Sondan) Üçüncü takımın forveti (oldukça) beğenildi sanıyorum.	I think that the forward player of the third team (from the bottom) was admired (much).
	EC	(Sondan) Üçüncü takımın forveti (oldukça) beğendiğini sanıyorum.	I think that the third team (from the bottom) admired the forward player (much).
12	**LC**	(Sağdan) Birinci kalecinin rakibi (sabah) dövüldü zannettik.	We considered that the rival of the first goalkeeper (from the right) was beaten (in the morning).
	EC	(Sağdan) Birinci kalecinin rakibi (sabah) dövdüğünü zannettik.	We considered that the first goalkeeper (from the right) beat the rival (in the morning).

Appendix B. (Continued)

		Experimental sentences	English translations
13	**LC**	(Siyah) Sakallı çiftçinin kızı (yazın) evlendirildi sanıyorum.	I think that the daughter of the farmer with (black) beard got married (in the summer).
	EC	(Siyah) Sakallı çiftçinin kızı (yazın) evlendirdiğini sanıyorum.	I think that the farmer with (black) beard married off the daughter (in the summer).
14	**LC**	(Toplam) Dokuz gazetecinin faili (suçlu) bulundu sanmıştım.	I thought that the murderer of nine journalists (in total) was found (guilty).
	EC	(Toplam) Dokuz gazetecinin faili (suçlu) bulduğunu sanmıştım.	I thought that nine journalists (in total) found the murderer (guilty).
15	**LC**	(Oldukça) Kalabalık cemaatin imamı (gece) dövüldü zannettik.	We considered that the priest of the (quite) crowded sect was beaten (at night).
	EC	(Oldukça) Kalabalık cemaatin imamı (gece) dövdüğünü zannettik.	We thought that the (quite) crowded sect beat the priest (at night).
16	**LC**	(Sağdan) Birinci yönetmenin kızı (sokakta) kaçırıldı sanıyorlar.	They think that the daughter of the first director (from the right) was kidnapped (on the street).
	EC	(Sağdan) Birinci yönetmenin kızı (sokakta) kaçırdığını sanıyorlar.	They think that the first director (from the right) kidnapped the girl (on the street).
17	**LC**	(Oldukça) Kalabalık aşiretin lideri (sokakta) vuruldu sanmışlar.	They thought that the leader of the (quite) crowded tribe was shot (in the street).
	EC	(Oldukça) Kalabalık aşiretin lideri (sokakta) vurduğunu sanmışlar.	They thought that the (quite) crowded tribe shot the leader (in the street).
18	**LC**	(Oldukça) Eski komşunun çocuğu (çok) sevildi zannettik.	We thought that the child of the (pretty) old (i.e., former) neighbor was liked (much).
	EC	(Oldukça) Eski komşunun çocuğu (çok) sevdiğini zannettik.	We thought that the (pretty) old (i.e., former) neighbor liked the child (much).
19	**LC**	(Uzun) Elbiseli kadının köpeği (sabah) yakalandı sanmıştım.	I thought that the dog of the woman with a (long) dress was caught (in the morning).
	EC	(Uzun) Elbiseli kadının köpeği (sabah) yakaladığını sanmıştım.	I thought that the woman with a (long) dress caught the dog (in the morning).
20	**LC**	(Oldukça) Minik bebeğin doktoru (yine) üzüldü sanıyorum.	I think that the doctor of the (very) little baby was upset (again).
	EC	(Oldukça) Minik bebeğin doktoru (yine) üzdüğünü sanıyorum.	I think that the (very) little baby upset the doctor (again).
21	**LC**	(Soldan) Birinci avukatın müvekkili (gece) bıçaklandı sanıyorlar.	They think that the client of the first lawyer (from the left) was stabbed (at night).
	EC	(Soldan) Birinci avukatın müvekkili (gece) bıçakladığını sanıyorlar.	They think that the first lawyer (from the left) stabbed the client (at night).
22	**LC**	(Yaklaşık) Otuz çocuğun danışmanı (biraz) yoruldu sandım.	I thought that the counselor of (nearly) thirty children was (a little) exhausted.
	EC	(Yaklaşık) Otuz çocuğun danışmanı (biraz) yorduğunu sandım.	I thought that (nearly) thirty children exhausted the counselor (a little).
23	**LC**	(Baştan) Üçüncü bakanın sekreteri sabah kovuldu sanıyoruz.	We think that the secretary of the third minister (from the beginning) was fired (in the morning).
	EC	(Baştan) Üçüncü bakanın sekreteri sabah kovduğunu sanıyoruz.	We think that the third minister (from the beginning) fired the secretary (in the morning).
24	**LC**	(Yirmi) İkinci şubenin müdürü (gece) sorgulandı zannettik.	We considered that the manager of the (twenty-) second branch was questioned (at night).
	EC	(Yirmi) İkinci şubenin müdürü (gece) sorguladığını zannettik.	We thought that the (twenty-) second branch questioned the manager (at night).