Timing of Syntactic and Rhythmic Effects on Ambiguity Resolution in Turkish: A Phoneme Restoration Study

Nazik Dinçtopal Deniz

(Department of Foreign Language Education,) Boğaziçi University, Turkey

Janet Dean Fodor

(Ph.D. Program in Linguistics,) The Graduate Center, The City University of New York, USA

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

Abstract

It has been shown that speakers use prosodic cues to disambiguate the syntactic structure of a sentence and listeners are sensitive to such cues. But the distribution of prosodic boundaries has been shown to depend on the lengths of constituents as well as the syntactic structure of utterances. Hence, it is possible that listeners are sensitive to these alternative reasons (i.e., syntactic or length-related) for why a speaker might introduce a prosodic break (Clifton, Carlson, & Frazier, 2006). The present study of Turkish employs a phoneme restoration paradigm to investigate more closely the time-course of three factors (prosodic cues, syntactic Late Closure, and phrase length effects) in the comprehension of a late/early closure ambiguity. The results confirm a significant role of prosody in restoring missing disambiguating phonemes; listeners tended to maintain an analysis that was syntactically or prosodically favored on-line. Notably, they did not generally revise that decision in face of the rhythmic factor of phrase lengths. This may be because length contrasts become fully apparent only later in a sentence. This is supported by the fact that when tested post-sententially in a previous study of Turkish (Deniz & Fodor, 2017), it was found that phrase lengths as well as prosodic and syntactic effects did influence parsing decisions, indicating that all three sources of guidance were at work. By comparing these two studies, with their different methodologies applied to the same materials, we document here the novel finding that rhythmic phrase length effects are indeed delayed, as previously contemplated by Clifton et al.

Introduction

Most of the early studies contributing to the development of models of human sentence processing relied on data from reading tasks, which exclude prosodic cues. However, as has been pointed out in more recent research, most language processing in everyday life occurs with spoken input, which does typically include rich prosodic information that could be helpful to the parser. It would be strange if this prosodic information were not accessible to the parser, because it is an integral part of the acoustic signal the parser receives.

Prosodic information can serve several functions, such as determining focus and topic, discriminating given and new information, disambiguating otherwise ambiguous structures, and expressing emotional states. Among other functions, the role of prosody as a disambiguator is of special interest to psycholinguistics since ambiguity resolution has long been used as a tool to shed light on the mechanisms underlying human sentence parsing.

It has been shown that speakers use prosody as a disambiguating cue in production (Lehiste, 1973; Price, Ostendorf, Shattuck-Hufnagel, & Fong, 1991; Schafer, Speer, Warren, & White, 2000; Snedeker & Trueswell, 2003). And listeners are sensitive to prosodic cues available in the speech input, not only in end-of-sentence judgments (Lehiste, 1973; Price et al., 1991; Schafer et al., 2000), but also at very early on-line stages of processing (Augurzky, 2006; Eckstein & Friederici, 2006; Kjelgaard & Speer, 1999; Marslen-Wilson, Tyler, Warren, Grenier, & Lee, 1992; Nagel, Shapiro, Tuller, & Nawy, 1996; Pauker, Itzhak, Baum, & Steinhauer, 2011; Speer, Kjelgaard, & Dobroth, 1996; Steinhauer, Alter, & Friederici, 1999; Steinhauer & Friederici, 2001; Stoyneshka, Fodor, & Fernández, 2010 among others; but cf., Watt & Murray, 1996).

Prosodic constituents usually correspond to constituents at some other level of linguistic analysis. In her Match theory Selkirk (2011) argues that the prosodic structure of an utterance largely matches its syntactic structure; i.e., the edges of syntactic constituents are matched to the edges of prosodic constituents. But certain phonological markedness constraints on prosodic structure may lead to violations of Match constraints and produce nonisomorphism between syntactic constituency and phonological domain structure. Although Turkish has been reported to show a high degree of mapping between syntactic structure and prosodic structure (e.g., İşsever, 2003), Güneş (2013) argues to the contrary, with data from syntactically isolated structures such as parentheticals (e.g., non-restrictive relative clauses) in Turkish, that “Turkish prosody mirrors the syntax and information structure to a lower degree of faithfulness than it has been assumed before”. If there were one-to-one mapping between syntax and prosody, the syntactically isolated structures would also be prosodically isolated as intonational phrases in Turkish. But the analyses revealed that prosodically they are not isolated. Their prosodic structure is similar to syntactically integrated structures (e.g., objects). She claims that “prosodic constraints may overweigh [syntax and information structure] in the formation and organization of prosodic structures.” (p. 3).

Turkish is a relatively understudied language in psycholinguistics. Studies on Turkish prosody and its role in sentence processing are even more limited. Given the diverging theoretical data on syntax-prosody mapping in Turkish (see above: İşsever, 2003; Güneş, 2013) experimental studies of the role of prosodic cues in disambiguating syntactic ambiguities in Turkish are of value.

Building upon an earlier study (Deniz & Fodor, 2017) we examine here the role of prosodic cues in processing a late/early closure temporary ambiguity in Turkish. The prior study (see further details below) showed that Turkish speakers use prosodic cues in their processing of syntactic ambiguities. The present study examines *when* such cues are used, and reveals that they are employed immediately. This study also investigates a different phenomenon, namely the role of phrase lengths in the perceived informativeness of prosodic boundaries in disambiguation of syntactic ambiguities, discussed here in terms of the Rational Speaker Hypothesis (Clifton, Carlson, & Frazier, 2002; Clifton et al., 2006).

Research on the syntax-prosody interface has shown that the production of prosodic boundaries in an utterance depends not only on its syntactic structure (Selkirk, 1984; Truckenbrodt, 1995) but also on the lengths of its constituents (e.g., Gee & Grosjean, 1983). In studies of perception it has been proposed (Clifton et al., 2002, 2006) that rhythmic constraints sensitive to phrase lengths are taken into account in assessing the significance of a prosodic boundary. This is the essence of the Rational Speaker Hypothesis (RSH) of Clifton and colleagues. The RSH maintains that listeners evaluate the informativeness of a prosodic break based on the length of the constituents it flanks, where prosodic breaks flanking short constituents are treated as more informative about syntax than breaks flanking long constituents, since only the latter would be motivated by optimal length considerations.

Clifton et al. (2006) investigated the RSH with Noun Phrase (NP) coordination and Adverb Phrase (AdvP) attachment ambiguities in English. Their experimental sentences included long vs. short NPs or long vs. short AdvPs. Listeners mostly adopted the prosodically appropriate interpretation for both short and long phrases, but did so significantly more often for the short phrases. Thus, as the RSH predicts, the listeners treated prosodic boundaries as more informative about the syntax when they flanked short constituents than when they flanked long constituents.

The RSH has further been supported with data from Turkish (Deniz & Fodor, 2017), a language typologically different from English. That study investigated RSH-induced rhythmic effects with a Turkish late/early closure ambiguity involving argument structure relations, which involved parsing a phrase as part of a subject NP or separately as an object NP (see example (1) below for further details of the ambiguity). The rhythmic factor of phrase lengths was pitted against the syntactically defined factor of Late Closure, to determine whether

Turkish listeners would be influenced by phrase lengths in accord with RSH or by the Late Closure strategy or possibly by both, in their interpretation of syntactically ambiguous sentences. The sentences were presented auditorily, with prosody that cooperated or conflicted with the morpho-syntactic structure or was neutral with respect to it. Participants' task was to listen to a sentence and indicate whether or not they had comprehended it (i.e., the "got it" task of Frazier, Clifton, & Randall, 1983). The results of that study showed an interplay between rhythmic effects as per the Rational Speaker hypothesis, and syntactic Late Closure, suggesting that phrase lengths modulate the parser's interpretation of prosodic boundaries alongside the syntactic Late Closure strategy, and that the two forces are not mutually exclusive. Further results of the study showed that the two factors affected the listeners' parsing decisions not only for cooperating prosody conditions but also for neutral prosody and conflicting prosody conditions.

Previous research has shown that prosodic boundary information is used on-line (Kjelgaard & Speer, 1999; Marslen-Wilson et al., 1992; Steinhauer, 2003, among others). Syntactic biases such as Late Closure have also been shown by previous research to affect parsing decisions rapidly (Frazier, 1978; Frazier & Rayner, 1982, among others). However, it is not clear how rapidly any rhythmic factors induced by prosody take effect or the extent to which they are the outcome of subsequent reflection; neither Clifton et al. nor Deniz and Fodor examined this. It is tested for Turkish in the experiment reported below, alongside the two other factors: prosodic cues and syntactic biases.

The experimental paradigm introduced in the present project is novel. It not only tests the time-course of prosodic and syntactic factors, for which there are alternative methods, but also makes it possible to examine the timing of rhythmic (phrase length) effects in processing spoken language input. The task is natural and not disruptive and is relatively easy to administer as it does not require advanced or expensive technology or exhaustive training. It is based on the phoneme restoration method previously used by Stoyneshka et al. (2010) in testing the role of prosody in disambiguating structural ambiguities in Bulgarian.

The phoneme restoration method

Stoyneshka et al. (2010) devised a new method to investigate the use of prosodic information in ambiguity resolution processes, namely the *phoneme restoration* paradigm. The phoneme restoration paradigm builds on the phoneme restoration effect first described by Warren (1970) in which a phoneme in a spoken sentence, although replaced by a cough, was perceived as intact. That is, when asked, listeners denied that any sound was missing; they thought they had heard a complete sentence with noise overlaid on it. Listeners could not even reliably identify where in the sentence the noise had occurred. The phoneme restoration effect has often been presented as a true perceptual illusion, but whether or not that is so, this method is welcome as a natural and unobtrusive method of gauging what analysis a listener is computing for a potentially ambiguous spoken input.

Stoyneshka et al. (op.cit.) employed this method in an investigation of prosodic influences on NP-coordination and relative clause (RC) attachment ambiguities in Bulgarian, inserting white noise in place of the phonemes that would have provided disambiguating information. Listeners' identification of those missing phonemes revealed which reading of the ambiguity they had computed. In one experiment, participants performed a post-sentence two-choice probe task, indicating which of two visually presented words they (thought they) had heard in the sentence. Results were highly sensitive to prosodic properties of the spoken sentences. Thus, phoneme restoration seems to be a promising technique for testing sensitivity to prosody in the current Turkish study.

Turkish is a language well-suited for this methodology because (like Bulgarian) its rich morphology allows us to replace just the disambiguating morphological information on the disambiguating verb, without masking the stem that carries its lexical meaning. Participants are told that the experiment concerns how well people can recognize words that are difficult to hear in noisy circumstances, as sometimes happens in everyday life due to poor transmission, such as on a cellphone. The aim is to draw listeners' attention away from the ambiguity in the materials.

In their Experiment 1, Stoyneshka et al. (2010) asked participants 'What word did you hear?' and participants chose from two words visually presented on a screen, one corresponding to each interpretation of the preceding spoken word string, which was disambiguated only by prosody. In the experiment we report here, a new variant of the phoneme restoration method was used. After hearing a sentence, a single word was visually presented and participants were asked 'Did you hear this word?', i.e., a simple one-word-probe task. Yes/no responses and RTs were recorded. This variant of the task (unlike the original two-word-probe task) was adopted on the assumption that it would be less likely to draw participants' attention to the existence of alternative possible morphological construals of the phoneme-replaced word.

The method employed in this study was also partially inspired by the cross-modal naming task which was previously used to measure on-line processing of prosodic boundaries (Kjelgaard & Speer, 1999; Marslen-Wilson et al., 1992; P. Warren, Grabe, & Nolan, 1995). In the cross-modal naming task, a fragment of the spoken sentence (up to a disambiguating word) is presented to the listeners and the time it takes for them to name (read aloud) a visually presented probe word which follows (either compatible with or conflicting with the prosodic structure of the spoken fragment) is measured. Although the cross-modal naming task can be informative about listeners' on-line parsing decisions, presenting only a fragment of the spoken utterance can interfere with the assignment or interpretation of natural prosodic breaks. It can also be disruptive of prosodic phrase lengths. The phoneme restoration task employed in the current study could therefore provide a better alternative, since it can reveal listeners' resolution of the ambiguity without interfering with the spoken sentence prosody and as soon as possible after the phrase length information becomes available.

Unlike the cross-modal naming task, in the phoneme restoration task employed here, the complete sentence is presented. Like cross-modal naming, the phoneme restoration task requires a response to the auditory stimulus without intrusion from comprehension questions or acceptability/grammaticality judgment tasks. The only intrusion is the visually presented probe word which appears immediately after the spoken sentence which is when the participants have full information regarding phrase lengths. In our materials the visual probe is presented only one-word (675 msec. on average) after the noise-containing word which is the penultimate word in the sentence. The final word (the matrix verb in (1a,b) below) was included because it was grammatically obligatory in early closure sentence versions as in (1b). It also played an important role in creating the illusion that there is no speech sound missing. Pre-tests had shown that if the noise replacement were sentence-final, participants were able to detect it as noise, so there was no illusion of speech masked by noise.

The present study

The Turkish ambiguity, its syntactic and prosodic properties, and phrase length manipulation

The study reported here is closely related to the previous study by Deniz and Fodor (henceforth D&F, 2017). The target sentences in that study and the present study included a temporary ambiguity which has not been a focus of prior psycholinguistic research. It concerns the attachment of a phrase either as the continuation of a subject NP or as standing alone as a following object NP. The ambiguity could be resolved either with late closure (LC) syntax as in (1a), or with early closure (EC) syntax as in (1b). (|| marks a prosodic boundary and / marks a syntactic boundary. Simplified syntactic bracketing illustrates the constituents relevant to the ambiguity.)

(1) a. LC: [$_{\text{CP}}$ Ø [$_{\text{CP}}$ [$_{\text{subject-NP}}$ (Yaklaşık) yedi öğrenci-nin psikoloğ-u] || /
Pro Nearly seven student-GEN psychologist-3SG.POSS-(NOM)
[$_{\text{embedded-VP}}$ (oldukça) sev-il-di]] san-ıyor-uz.]
much like-PASS-PAST think-PROG-1PL
'We think that the psychologist of (nearly) seven students was (much) liked.'

b. EC: [$_{\text{CP}}$ Ø [$_{\text{CP}}$ [$_{\text{subject-NP}}$ (Yaklaşık) yedi öğrenci-nin] || / [$_{\text{object-NP}}$ psikoloğ-u]
Pro Nearly seven student-GEN psychologist-ACC
[$_{\text{embedded-VP}}$ (oldukça) sev-diğ-i-ni]] san-ıyor-uz.]
(much) like-FN-3SG.POSS-ACC think-PROG-1PL
'We think that (nearly) seven students liked the psychologist (much).'

Both sentence types have a two-clause structure in which the final verb is in the matrix clause, and its subject is the initial *pro* in all cases. The embedded clauses in (1a,b) are either in the form of a tensed embedded clause as in (1a), or in the form of a nominalization as in (1b) (Erguvanlı, 1984). Tensed embedded clauses such as (1a) are very similar to main clauses as they receive regular nominative case (which is null) on the subject and standard verbal endings. In nominalized embedded clauses such as (1b), the whole embedded clause is in the form of a genitive-possessive noun phrase (NP) inside which the subject NP is marked with genitive case and the embedded verb is marked with a possessive suffix after being nominalized. (1b) is thus incomplete without the matrix verb. Although the subordinate clause in (1a) could be processed as an independent clause if the final verb were missing, to keep the experimental sentences structurally consistent, we used a two-clause structure in (1a) as well.

The temporary ambiguity is due to the homophony of the morphological marking (*-u*) of the second noun (*psychologist*), which can be processed either as a third person singular possessive marker as in (1a) or as an accusative case marker as in (1b). In (1a), the first and second noun (*student* and *psychologist*) together form an NP which functions as the nominative subject of the passive verb *sev-il-di* in the subordinate clause. This amounts to late closure of the subject: *student-GEN psychologist-3SG.POSS*. By contrast, in (1b), only the first noun constitutes the subordinate clause subject (genitive-marked). The subject is closed early, ending at *student-GEN*, and then *psikoloğ-u* must be interpreted as the accusative object (*psychologist-ACC*) because the subordinate verb (*sev-diğ-i-ni* in (1b)) is active and obligatorily transitive. The disambiguating verbs were matched for phonological length and corpus frequency in the experimental materials.

According to the Late Closure principle, incoming words are preferentially attached into the phrase currently being processed (Frazier, 1987, p. 562). Thus, when Late Closure is at work (setting prosody aside for the moment), it is predicted that (1a) will be easier to process than the matched (1b) because after processing the first NP, the parser will attach the incoming item to the phrase currently being processed and form a complex NP, *öğrenci-nin psikoloğ-u* (student's psychologist), interpreting the morpheme *–u* as the third person singular possessive marker. This will be appropriate for the interpretation shown in (1a), with a passive verb, where the complex NP functions as the subject of the sentence and there is no object. For (1b), however, the parser would discover that this analysis was incorrect when it subsequently encounters the transitive verb which needs an accusative case-marked direct object. Thus, the second NP must be re-processed as the direct object of the embedded verb and the first NP must be re-processed as the subject of the embedded clause.

Note that the Minimal Attachment parsing strategy, also syntactic in origin, is also relevant for the parsing of the temporary ambiguity in (1). The Minimal Attachment strategy prevents postulation of any potentially unnecessary nodes (Frazier, 1987). On encountering NP-GEN (i.e., before seeing/hearing the second NP), it is simpler for the parser to build NP-GEN as a modifier inside a complex NP subject than to build it as the subject of a subordinate clause, which would require a more complex sentence structure. Thus, Late Closure and Minimal Attachment concur that (1a) is preferable to (1b).

Prosodic phrasing provides a clear disambiguating cue to the correct parse of the first few words of the sentence. The default topic position in Turkish is sentence-initial. Since Turkish is an SOV language, subjects typically appear at the beginning of sentences and are therefore the default or natural topic unless a different word order is used (Erguvanlı, 1984; Göksel & Kerslake, 2005; İşsever, 2003; Kornfilt, 2011). Erguvanlı (1984) mentions that it is also possible to topicalize a constituent of an embedded clause (including nominalized or finite noun clauses) by either moving it out of the clause or by placing a prosodic boundary after it in situ. In all the examples above, it was the subject phrase of the subordinate clause that was topicalized. This resulted in an intonational phrase (IPh) boundary after the subject of the subordinate clause in each of (1a) and (1b).

A topic in Turkish is associated prosodically with an exaggerated high pitch accent, realized as a rising tone, with an optional pause following it (Kamali, 2008; Vallduví & Engdahl, 1996). These acoustic correlates are associated with IPh boundaries by Kan (2009). Thus, a topic in Turkish, regardless of its length, constitutes a separate IPh, separated from following words by a strongly marked IPh boundary. The experimental sentences in the present study were uttered with an IPh boundary after the subjects. The location of that IPh boundary in utterances of sentences like (1a) and (1b) provides a clear prosodic cue favoring their respective syntactic structures, eliminating the ambiguity before the morphosyntactic disambiguation at the embedded verb.

All target items consisted of six prosodic words (PWds). To manipulate phrase length variations in the experimental materials, modifiers as shown in parentheses in (1a,b) above were used to lengthen either the subject or the verb phrase (but not both) of the subordinate clause. In both disambiguating prosody conditions, the 6 PWds were pronounced as two prosodic phrases, but the groupings differed between 4+2 PWds and 3+3 PWds in the lengthened subject condition and between 3+3 PWds and 2+4 PWds, in the lengthened VP condition. All these patterns of phrase lengths are fully acceptable in Turkish though in light of data from Nash (1973), prosodic phrases consisting of 2 PWds would be perceived as short in Turkish, while those with 3 or 4 PWds would fall within the typical range. Sentence versions without any prosodic boundary were also tested; see below.

The corpus data and the norming study

The experimental sentences had been selected on the basis of a norming study to control for any potential inherent bias for LC or EC items. Noun phrase sequences which had a potentially ambiguous morphological ending on the second NP (3SG.POSS or ACC; i.e., LC or EC interpretation respectively) 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, were selected to be used in creating a pool of potential experimental sentences. Forty sentences (with LC and EC continuations following 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 an appropriate adjective phrase (e.g., *(yaklaşık) yedi*) and verb phrase (e.g., *(oldukça) sevildi/sevdiğini*) respectively; see examples (1a,b) above.

A norming study following the method in Webman-Shafran and Fodor (2016) was used to pretest these potential experimental sentences to eliminate ones that were strongly biased towards one interpretation. The two 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.

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: $t(23) = .923$, $p = .366$. 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.

Materials

Materials design is crucial to the interpretation of experimental results to be reported here, and the design will therefore be presented below in some detail.

For purpose of comparison with earlier findings the sentences had identical syntactic and prosodic features and phrase length distributions as in D&F (2017) but the disambiguating morpho-syntactic information in the sentence was replaced with pink noise in this experiment (shown as '' in the examples below). Following the lead of Bashford et al. (1992) and Bashford and Warren (1987), pink noise (which has equal power per octave) was employed instead of white noise, a small divergence from Stoyneshka et al. (2010).

Each sentence recording was followed by a single visual probe word whose morphology was compatible only with an LC interpretation of the sentence, or only with an EC interpretation. Participants could accept or reject the word as having been present in the sentence, presumably on the basis of how they had parsed the sentence on-line and thus what phonemes they thought they had heard. The only disambiguating cue to the syntactic analysis was the prosodic phrasing, in those sentences in which disambiguating prosody was present. Note that for baseline purposes there were also sentences which did not have any disambiguating prosody (see discussion below).

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 (lengthened subject vs. lengthened VP) as well as their prosody and syntactic structure. The complexity of the design made it impractical to test the length contrast within subjects, 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 morphological structure of the probe word were manipulated within subjects.

There were six presentation conditions in each length manipulation, as illustrated in (2) and (3). Detailed glosses of these sentences are given in (1a,b) above. The glosses are identical across all versions in (2) and (3); they are omitted here to highlight the prosodic properties and length manipulations in the spoken sentences. As before, the symbol || is used to mark prosodic boundaries. Also, the word shown in bold in the neutral prosody versions (2c) and (3c) received an H* pitch accent indicating contrastive focus in order to reduce prosodic variation in the remainder of the sentence.

(2) Lengthened Subject

Spoken Sentence	*Visual Probe* LC /EC
a. LC Prosody Yaklaşık yedi öğrencinin psikoloğu || sevi sanıyoruz.	sevildi / sevdiğini
b. EC Prosody Yaklaşık yedi öğrencinin || psikoloğu sevi sanıyoruz.	sevildi / sevdiğini
c. Neutral Prosody Yaklaşık **yedi** öğrencinin psikoloğu sevi sanıyoruz.	sevildi / sevdiğini

LC Syntax: 'We think that the psychologist of nearly seven students was liked.'

EC Syntax: 'We think that nearly seven students liked the psychologist.'

(3) Lengthened VP

Spoken Sentence	*Visual Probe* LC /EC
a. LC Prosody Yedi öğrencinin psikoloğu || oldukça sevi sanıyoruz.	sevildi / sevdiğini
b. EC Prosody Yedi öğrencinin || psikoloğu oldukça sevi sanıyoruz.	sevildi / sevdiğini
c. Neutral Prosody **Yedi** öğrencinin psikoloğu oldukça sevi sanıyoruz.	sevildi / sevdiğini

LC Syntax: 'We think that the psychologist of seven students was much liked.'

EC Syntax: 'We think that seven students liked the psychologist much.'

As noted above, the LC and EC prosody sentences had an IPh boundary after their subjects in accord with the topicalization facts in Turkish, while sentences with neutral prosody had no prosodic boundary in the corresponding region of the sentence. In order for this neutral prosody to sound natural, the modifier of the sentence-initial phrase (the subject of the sentence) received a contrastive accent, which tends to mute prosodic variation in the remainder of the sentence. This strategy is similar to the one used by Kjelgaard and Speer (1999), who placed contrastive focus on the subject of a temporarily ambiguous first clause. Note that this baseline prosody is not argued to be neutral in any absolute sense. Rather, it serves as a baseline by not providing any prosodic disambiguating cue. Figures 1-4 illustrate the LC, EC and neutral prosody conditions, in the lengthened subject condition, with disambiguating morphology of an overt embedded verb.

[insert Figures 1-4.]

Note that in the lengthened subject condition as in (2), LC prosody (2a) yielded an unbalanced prosodic phrasing consisting of 4+2 PWds. The LC probe would thus be congruent with the 4+2 PWds sentence prosody, while the EC probe would be incongruent with that sentence prosody. By contrast, EC prosody in the lengthened subject condition (2b) would yield balanced lengths (3+3 PWds). The EC probe would therefore be congruent with the 3+3 PWds sentence prosody, whereas the LC probe would be incongruent with that sentence prosody. The neutral prosody condition (2c) did not have any 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, as noted above. The participant’s response to the visual probe would show how they had construed the sentence structure. 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 whatever syntactic phrasing they had happened to compute for the sentence. Either probe word would be compatible with one prosody of the spoken sentence, but the participant's response to the visual probe word would presumably be influenced by, and thus would reveal, the syntactic structure that they had computed for the sentence.

In the lengthened VP condition as in (3), prosodic lengths were reversed compared with the lengthened subject condition: It was LC prosody (3a) that yielded balanced phrasing (3+3 PWds). The LC probe would be congruent with the 3+3 PWds sentence prosody, whereas the EC probe would be incongruent with the sentence prosody. The EC prosody in the lengthened VP condition (3b) yielded unbalanced phrasing (2+4 PWds). The EC probe would be congruent with the 2+4 PWds sentence prosody, whereas the LC probe would be incongruent with it. The neutral prosody condition (3c) did not have any prosodic boundary signaling either LC or EC structure so a hearer might impose either a symmetric 3+3 syntactic division which would be consistent with LC morphology, or a 2+4 division which would be consistent with EC morphology. (Which interpretation listeners had favored could be revealed by RT data.)

It is important to note that in the incongruent probe conditions the mismatch between the probe and the prosody of the sentence might lead to the probe being rejected. Alternatively, it is possible that a participant would take the probe word as a *prompt,* indicating what to search for in the sentence, and accept it. In that case, an EC probe after the LC prosody sentence might cause the LC prosody to be ignored, in favor of a syntactic EC structure, i.e., 3+3 PWds phrasing in the lengthened subject condition and 2+4 phrasing in the lengthened VP condition. Likewise, an LC probe after an EC prosody sentence, if interpreted as a prompt, overriding the prosody, would imply an LC structure, i.e., 4+2 PWds in the lengthened subject condition and 3+3 PWds in the lengthened VP condition.

Table 1 summarizes the phrase lengths in the between-subjects conditions (lengthened subject, lengthened VP) and the within-subjects conditions (the sentence prosody: LC, EC and neutral, and the congruency of the visual probe: congruent and incongruent). Note that it is not possible to predict whether an LC or EC visual probe will be compatible with the syntactic structure computed in the neutral prosody condition.

Table 1. Phrase lengths in each condition in the study. Congruent probe conditions are in bold face. Incongruent probe conditions are italicized.

Between-Subjects Conditions	**Within-Subjects Conditions**		
		LC Probe	EC Probe
Lengthened Subject	LC Prosody (4+2)	**4+2 PWds**	*3+3 PWds*
	EC Prosody (3+3)	*4+2 PWds*	**3+3 PWds**
	Neutral Prosody (6 PWds)	4+2 PWds	3+3 PWds
Lengthened VP	LC Prosody (3+3)	**3+3 PWds**	*2+4 PWds*
	EC Prosody (2+4)	*3+3 PWds*	**2+4 PWds**
	Neutral Prosody (6 PWds)	3+3 PWds	2+4 PWds

Care was taken to create an acoustically neutral context for both LC and EC morphology conditions. It is known that a speech sound in continuous speech is influenced by its preceding or following sounds, and co-articulatory cues have been reported to cross even syllable boundaries (Magen, 1997). Thus, even the sounds which are not immediately adjacent to the disambiguating phonemes could potentially carry some phonetic information about them. In order to prevent co-articulation effects, the spoken sentences in the experiment were originally uttered and recorded with the glottal fricative /h/ in place of the disambiguating morphemes. The glottal fricative /h/ is produced at the glottis and provides a neutral context, free of supralaryngeal articulation (Tunley, 1999). Warren and Sherman (1974) used a similar strategy for English to show that phoneme restoration succeeds even when co-articulation is controlled for (details can be found in R. Warren & Sherman, 1974, p. 152). In the current study, the co-articulation-neutral /h/ sound was used in all target items (see details in Figure 5) and in some proportion of the fillers. That this achieved its purpose is confirmed by the prosodic and syntactic effects in the main experiments. (See below.)

The disambiguating phonemes' average length (averaged across LC morphology and EC morphology) was used to determine the length of the inserted noise for each item in the present study. For disambiguating prosody sentences, it was computed by averaging the length of LC/EC phonemes prior to /h/ substitution. Figure 5 exemplifies how the splicing, length averaging and noise replacement was done.

[insert Figure 5.]

Following Bashford et al. (1992) the noise was presented at 80 dB, which was approximately 15 dB louder than the average amplitude of the rest of the utterance. The reason for having the noise louder than the actual speech signal was to contribute to the illusion that the missing information was indeed masked by the loud noise, not replaced by it. Pink noise creation and acoustic manipulation of the sentential items were all done using the Praat software (Boersma & Weenink, 2009).

The maximum duration of noise permitting reliable perceptual restoration of the missing phoneme has been reported (for English) to be 333 msec. in Powers and Wilcox (1977) and 304 msec. in Bashford and Warren (1987). The duration of the noise in the Turkish experimental items had a range from 184 msec. to 262 msec. (mean = 221 msec.) falling well below the ceiling for English. Thus, the noise duration in the current experiments should allow for successful restoration.

A prosodic boundary was acoustically characterized by a pause following the relevant NP, and increased pitch on and lengthening of the pre-boundary word. Acoustic analyses were conducted to ensure that all items had the intended prosodic properties. Table 2 shows the average maximum F0 and duration for NP1s (e.g., *öğrencinin* '*student*') and for NP2s (e.g., *psikoloğu* '*psychologist*') in (2) and (3), and the duration of pauses (where applicable). *t*-tests confirmed that NP1 and NP2 durations and maximum F0's differed significantly in the intended directions for LC and EC prosody conditions. *t*-statistics are given in Table 2 beneath the cells that were compared.

Table 2. Acoustic properties of experimental items in the study. *df* = 23 in all comparisons and *p*'s < .001.

Between subject conditions	**Within subject conditions**					
	Prosody	**F0 maximum (Hz)**		**Duration (ms)**		**Pause (ms)**
		NP1	*NP2*	*NP1*	*NP2*	
Lengthened Subject	*LC*	378	468	479	611	360 (after NP2)
	EC	476	267	662	413	393 (after NP1)
		$t = 8.89$	$t = -46.67$	$t = 30.66$	$t = -24.49$	
	Neutral	271	215	453	381	NA
Lengthened VP	*LC*	416	475	493	663	306 (after NP2)
	EC	474	271	667	403	367 (after NP1)
		$t = 8.18$	$t = -34.97$	$t = 22.68$	$t = -24.79$	
	Neutral	306	223	463	362	NA

As can be observed from the table, the pre-boundary words (NP1 for EC and NP2 for LC) were associated with higher F0 values and longer durations in both length conditions, as predicted. They were also followed by a pause. These cues were absent in the neutral prosody condition. Thus, the acoustic analyses confirm that the sentential items had properties appropriately associated with LC and EC prosody and neutral prosody.

There were 24 experimental item sets, each set in twelve conditions manipulating length (lengthened subject, lengthened VP) and prosody of the spoken sentence (LC, EC and neutral) and the morphology of the probe word (LC, EC). Length was a between-subjects condition. Each length condition had six lists counterbalancing for the three prosody types and the two probes. Each list included 24 target sentences. These were intermingled with 72 fillers, of which 24 belonged to a different syntactic ambiguity. The remaining 48 fillers were unambiguous and consisted of 24 control items with active/passive morphology (details below) and 24 fillers with varied structures. In addition, there were 10 unambiguous practice items prior to the beginning of the experiment, and 10 implicit ‘warm-up’ items, 5 at the beginning of each list and 5 half way through, before which participants were encouraged to take a rest break. Thus, each list following the practice session consisted of 106 sentences.

It was anticipated that there would be a potential unintended bias for the EC interpretation of the spoken sentence since it would contain an active form of the verb (or equivalently a bias for the active form of the visual probe word), as opposed to the LC interpretation under which the verb would be in passive form. Some piloting prior to the main phoneme restoration experiments had indicated this. In order to control for this confound, additional filler items were created. These control filler items had active/passive syntactic structures similar to the experimental items but they were not ambiguous. The ambiguities tested in this study occurred because the second NP ended in a consonant (e.g., *psikolog* ‘psychologist’ in (2) and (3)). For a word with a final consonant the usual case morphology is reduced by deletion of its initial consonant (as in example (1) above, in which the consonant *-s* of the morpheme *-sI* for 3SG.POSS and the consonant *-y* of the morpheme *–yI* for ACC are deleted on *psikolog*). This is what creates the ambiguity in the experimental sentences. In the unambiguous filler sentences the stem of NP2 ended in a vowel preventing any consonant deletion in the case morpheme, which eliminated the morphological ambiguity of the NP2. Thus, the control filler sentences had the same syntactic structure as the experimental sentences but with unambiguous morphology. These unambiguous control fillers were uttered with neutral prosody to ensure that any preference in interpretation must be due solely to a lexical, not a prosodic, preference. As with the target sentences, the post-sentence visual probes for these fillers were either congruent or incongruent with the morpho-syntactic information in the sentence. Responses and response times to these control filler items would serve to balance the anticipated bias for an active verb in the target items.

All fillers included noise replacing a few phonemes in the penultimate or final word in the sentence.

Participants

The experiment was conducted in Turkey. 96 native speakers of Turkish took part in the experiment; all were college students. 48 of those participants were in the lengthened subject condition (mean age = 21.4, 5 males) and 48 were in the lengthened VP condition (mean age = 21.1; 4 males). Participants received 15 Turkish liras (~$8.5 at the time of the experiment) for their participation.

Procedure

Participants were tested individually, seated in front of a computer screen in a quiet room. They were presented with the spoken sentences via noise-cancelling head phones. At the end of each spoken sentence, a visual probe word appeared on the screen. Participants' task was to listen to the sentence and indicate, by pressing the *yes* or *no* button marked on the keyboard, whether they had heard the visual probe word in the spoken sentence. The first author stayed with each participant during a practice session and answered any questions the participants might have. The task took 15-20 minutes to complete after instruction and practice.

Yes/no responses were taken as indicating which verb form the listeners thought they had heard in the sentence; in case that did not match the probe word, the listener could respond *no*. But it is possible, as noted above, that participants might sometimes interpret the visually presented word as a prompt, rather than a probe. That is, they might think that it was a hint as to which word had been in the sentence, and that could lead them to reanalyze the sentence, at some cost in the response time. Therefore, response times were recorded and analyzed as a potential indication of perceived confirmation or conflict, and possible re-analysis (see discussion below in the Results Section).

Predictions

The extensive research on prosody in sentence processing (Lehiste, 1973 and since) has shown that listeners are sensitive to prosodic cues in disambiguating the syntactic structure of an utterance. Although Turkish has not been much investigated in this regard, it is reasonable to anticipate that Turkish speakers will also be responsive to the prosody of the spoken sentence in the present study. The LC or EC prosody will inform them in their computation of the syntax of the structure and restoration of missing phonemes in the input. Thus, we predict that the participants' *yes* responses to the probe word will be most frequent for the conditions where the probe word is congruent with the prosody of the sentence and least frequent for the incongruent probes, with prosody-neutral probes in between.

Previous research on other constructions provides evidence that prosodic cues are used rapidly (Kjelgaard & Speer, 1999; Marslen-Wilson et al., 1992; Nagel et al., 1996 among others). If that is also true for the Turkish ambiguity investigated here, it is predicted that the participants' restoration of the missing phonemes would be in line with the prosody of the spoken sentence that they had listened to, with *yes* response times to the probe word being the shortest for the prosody-congruent probes, longest for the prosody-incongruent probes and intermediate for prosody-neutral probes. For conditions where the prosody was neutral and did not bias for either LC or EC attachment, an LC bias is also predicted (the Late Closure strategy: Frazier, 1978; and extensive other findings).

Evidence that both rhythmic and syntactic pressures are operative in processing the LC/EC ambiguity examined in the present study is presented in D&F (2017). But no prior study has provided evidence on the timing of RSH effects, though this question was raised by Clifton et al. (2006) as noted above. If phrase length information as per the RSH is also used without delay, the results of the present experiment will be similar to those of D&F. In that case, there is a clear prediction that the shortest response times (RTs) should obtain wherever

Late Closure and RSH both cooperate with the correct syntactic structure. An instance of this would be the 4+2 LC items in the lengthened subject condition (e.g., (2a) congruent probe above). The longest RTs should be observed wherever the two influences conspire against the correct syntactic structure (e.g., 3+3 EC items in the lengthened subject condition as in (2b) congruent probe). Where the two oppose each other (as in both LC items and EC items in the lengthened VP condition; e.g., (3a,b) congruent probe), intermediate RTs may be observed.

On the other hand, if RSH effects are less immediate than LC effects and if the phoneme restoration method is sensitive to subtle timing differences, then, as per the Late Closure strategy, a more immediate response to LC probes than to EC probes would be observed in both length conditions.

Data Analysis and Results

Yes/no response data and RTs were analyzed using the R statistical computing software, version 2.15.2 (R Core Team, 2012). The analyses controlled for participants' responding faster as the experiment progressed (β = -1.76, *SE* = .38, *t* = -4.61, $p < .001$) and for their bias for the active form of the probe word (response data odds ratio: β = 2.98, *SE* = .06, *z* = 5.42, $p < .001$; RT data: β = 51.17, *SE* = 6.17, *t* = 9.96, $p < .001$). The RT data will be reported below, following presentation of the response data. The *p*-values for the predictors in linear models were calculated using Satterthwaite approximations for degrees of freedom (Luke, 2017), which was available in the *lmerTest* package. The *p*-values for pairwise comparisons and logistic regressions were elicited automatically through the relevant packages used.

<u>Response data.</u> The analyses of acceptance data were run via a mixed effects logistic regression model using the *glmer* function in the *lme4* package. Phrase lengths (lengthened subject, lengthened VP), prosodic congruency (congruent, incongruent and neutral) and the morpho-syntax of the visual probe (LC, EC) were entered as fixed factors, and subjects and items were random factors. The acceptance data were in the predicted direction. Participants

were more likely to respond *yes* to the prosody-congruent probe, followed by the prosody-neutral and then the prosody-incongruent probes. This was so for both the general data (odds ratios for congruent vs. neutral: $\beta = 2.07$, $SE = .02$, $z = 5.37$, $p < .001$; incongruent vs. neutral: $\beta = .32$, $SE = .03$, $z = -8.37$, $p < .001$) and for specific length conditions: lengthened subject condition, congruent vs. neutral: $\beta = 2.38$, $SE = 1.22$, $z = 4.31$, $p < .001$; incongruent vs. neutral: $\beta = .24$, $SE = 1.21$, $z = -7.11$, $p < .001$; lengthened VP condition: prosody-congruent vs. prosody-neutral: $\beta = 1.92$, $SE = 1.19$, $z = 3.56$, $p < .001$; prosody-incongruent vs. prosody-neutral: $\beta = .39$, $SE = 1.2$ $z = -5.11$, $p < .001$). The models allowed for random intercepts and random slopes for the active bias predictor for both subjects and items (χ^2s > 78.72, *p*'s < .001). There was no reliable difference in *yes* responses attributable to any factor other than prosodic congruency of the probe, neither an overall LC probe/EC probe difference nor a phrase length effect (*p*'s > .05). Using a log-likelihood ratio test, the simple models with each predictor variable were compared to the more complex models that involved both predictors, but the complex models were not significantly better than the simple model with prosody as the predictor. There was also no interaction involving these variables (χ^2s < 0.6, *p*'s > .5).

Figure 6 below shows the general distribution of the *yes* responses[1]. Data for specific length conditions (lengthened subject, lengthened VP) mirrored this general pattern.

[insert Figure 6.]

As the analyses and Figure 6 show, the participants were strongly influenced by prosodic information in restoring missing phonemes in these sentences. Participants most often judged the probe word as having been in the sentence they had heard when it corresponded to a prosody-congruent word in the spoken sentence; they rejected the probe word most often when it was not compatible with the prosody of the sentence.

[1]The graph here shows the *adjusted* probability of *yes* responses (rather than actual probabilities) based on statistical analyses controlling for bias for the active form of the visual probe.

Methodologically, these data confirm that the single word probe recognition method is sensitive to sentential prosody, similar to the two-choice probe method that Stoyneshka et al. (2010) employed in their study of coordination and relative clause constructions in Bulgarian. This is a welcome finding in as much as the single-word task reduces the listener's processing load and thus may give finer estimates of timing.

RT data. Inspection of the RT data showed divergence from normality, $W = .80$, $p < .001$, $D = .15$, $p < .001$. Hence, the RT data were log-transformed (Baayen & Milin, 2010; Ratcliff, 1993) and any outliers that were above/below ±1.5 x interquartile range were excluded from the log-transformed data (approximately 5% of the data). The RTs were then entered into a mixed effects model using the *lmer* function in the *lme4* package. Length (lengthened subject and lengthened VP), congruency (congruent, incongruent, neutral) and morpho-syntax of the probe (LC, EC) were fit as fixed factors, and subjects and items as random factors.

In this probe task the data were collected in two forms: *yes/no* responses and RTs. The RTs were recorded separately for the *yes* and the *no* responses, since they could contribute to the RT data differently. Whereas *yes* responses could yield shorter RTs for congruent probes (and perhaps longer RTs for incongruent probes), *no* responses could yield shorter RTs for incongruent probes (and perhaps longer RTs for congruent probes). Thus, the *yes/no* responses to the experimental items were entered into the model to check whether or not they affected the RTs. The analysis indicated that the response *yes* or *no* did influence the RT. In general, *yes* responses were faster than *no* responses ($\beta = -149.43$, $SE = 22.68$, $t = -6.28$, $p < .001$). In the main analyses, this variable also showed a strong interaction with the probe's congruency with the sentence prosody, $\chi^2(4) = 61.33$, $p < .001$[2]. The pairwise comparisons indicated that *yes* and *no* responses showed patterns that were in mutually opposite directions

[2] In previous research some researchers conducted response time analyses by combining *yes* and *no* responses (e.g., Kjelgaard and Speer (1999) in their phonosyntactic grammaticality judgment task). However, for the method used in this study, it seemed appropriate to analyze *yes* and *no* responses separately.

(e.g., faster *yes* and slower *no* responses for a probe congruent with the prosody, and faster *no* and slower *yes* responses for a probe incongruent with the prosody, *t'*s > 2.09). Here, we will only report the analyses conducted for the RTs for *yes* responses.

The analyses of RTs for *yes* responses did not show any main effect of phrase length (β = 42.55, *SE* = 58.65, *t* = .73, *p* > .05) but the model which included an interaction of length (lengthened subject, lengthened VP), prosodic congruency (congruent, incongruent and neutral) and morpho-syntax of the visual probe (LC, EC) was significantly better than a less complex model where only congruency and morpho-syntax interacted, $\chi^2(6) = 13.97$, $p < .05$.

The following section presents separate analyses for the lengthened subject and lengthened VP conditions, to better understand the nature of this interaction. Note that the models (simple or complex) for both the general data and the specific length conditions allowed only for random intercepts. The models with random slopes for the variables were not significantly better than random-intercepts-only models, χ^2s < 6.21, *p*'s > .3.

Lengthened subject condition. The analyses for *yes* responses were run on prosodic congruency of the probe word (congruent, incongruent, and neutral) and morpho-syntax of the probe (LC and EC), and their interaction. The model for congruency suggested no reliable difference in RTs for *yes* responses to probes that were not prosodically biased (i.e., neutral prosody) and probes congruent with the prosody of the sentence, $\beta = 81.68$, *SE* = 45.3, *t* = 1.9, *p* = .06. Analyses on congruency also indicated that prosody-neutral probes were processed faster than prosody-incongruent probes, $\beta = 285.05$, *SE* = 62.18, *t* = 5.24, *p* < .001. The model with morpho-syntax of the probe as a predictor variable showed that there was not a significant difference between the LC and EC probes, $\beta = -59.35$, *SE* = 51.52, *t* = -1.15, *p* = .25.

However, analyses showed an interaction of the two predictors: $\chi^2(3) = 12.07$, $p < .005$ for the comparison of the simple model for congruency vs. a complex model for congruency and morpho-syntax and $\chi^2(4) = 37.78$, $p < .001$ for the comparison of the simple model for morpho-syntax vs. a complex model for congruency and morpho-syntax.

During model criticism, data points with standardized residuals below/above 2.5 standard deviations were excluded from the analyses as well as overly influential subjects (N = 2), items (N = 1) and individual data points (N = 2).

RTs from the remaining participants and items can be observed in Figure 7[3].

[insert Figure 7.]

Planned pairwise comparisons using the *glht* function showed that when the probe was congruent with the prosody, there was no reliable RT difference between the LC probe and the EC probe ($\beta = -85.92$, $SE = 66.51$, $z = -1.28$, $p = .2$). *Yes* responses to EC probes were faster than to LC probes when the probe was incongruent with the prosody ($\beta = 326.66$, $SE = 159.53$, $z = 2.43$, $p < .05$). Responses to LC probes were faster than to EC probes when the prosody was neutral ($\beta = -204.95$, $SE = 69.9$, $z = -2.75$, $p < .01$). These results will be discussed after presentation of analyses for the lengthened VP condition.

Lengthened VP condition. As in the lengthened subject condition the main analyses were run on the prosodic congruency (congruent, incongruent, neutral) and the morpho-syntax of the visual probe (LC, EC), and their interaction. The model for prosodic congruency showed no reliable RT difference between prosody-congruent probes and prosody-neutral probes, $\beta = -46.36$, $SE = 47.7$, $t = -.97$, $p = .33$. Prosody-neutral probes were processed reliably faster than prosody-incongruent probes, $\beta = 226.47$, $SE = 67.45$, $t = 3.73$, $p < .001$. The model with morpho-syntax of the visual probe as a predictor variable showed that overall there was not a significant difference between the LC and EC visual probe, $\beta = -77.71$, $SE = 60.82$, $t = -1.27$, $p = .21$.

[3] The graph here and the one in Figure 8 show *adjusted* RTs (rather than actual RTs) based on statistical analyses controlling for bias for active form of the visual probe (see above).

The analyses also indicated that a model including an interaction of the two predictors was significantly better than at least one of the simpler models: $\chi^2(3) = 2.21$, $p = .48$ for the comparison of the simple model for congruency vs. a complex model for congruency and morpho-syntax; $\chi^2(4) = 23.91$, $p < .001$ for the comparison of the simple model for morpho-syntax vs. a complex model for congruency and morpho-syntax. Thus, the model with the interaction was kept.

During model criticism, data points with standardized residuals below/above 2.5 standard deviations were excluded from the analyses as well as overly influential subjects (N = 4) and data points (N = 2). There were no overly influential items. RTs from the remaining participants can be observed in Figure 8.

[insert Figure 8.]

Planned pairwise comparisons showed that in the prosody-congruent condition, there was no reliable difference between the LC and EC probes (β = -29.61, *SE* = 83.93, z = -.36, p = .7). There was a reliable LC advantage in the prosody-incongruent condition (β = -278.3, *SE* = 94.12, z = -2.73, $p < .01$) and the prosody-neutral condition (β = -184.35, *SE* = 83.53, z = -2.11, $p < .05$).

The data show the pattern in Table 3, in which two of the three conditions across the two length manipulations (lengthened subject, lengthened VP) give comparable results. The only case in which the results for the two length conditions differ is where the probe is incongruent with the prosody. This and the other results presented above will be addressed in the Discussion Section.

Table 3. Lengthened subject and lengthened VP conditions, summary of *yes* response time data pattern. ‘<’ indicates faster processing, ‘=’ indicates no significant difference in processing time. All inequalities in the table are confirmed at $p < .05$ or smaller.

	Probe **Prosody-congruent**	**Prosody-neutral**	**Prosody-incongruent**
Lengthened subject	LC = EC	LC < EC	LC > EC
Lengthened VP	LC = EC	LC < EC	LC < EC

Discussion

Summarizing the above results, we consider first the *yes/no* data, which suggest that listeners are sensitive to the prosodic contour of the sentence in restoring the phonemes that are missing in the spoken input. This is reflected in their *yes* responses to the visual probe, where they were most likely to accept the visual probe when it was supportive of the prosody of the sentence they had listened to and presumably with the syntactic analysis they had computed on the basis of that prosody. This suggests that they restored the phoneme that would be compatible with the prosodic contour of the sentence, thus contributing to the high rates of *yes* responses to congruent probes, and the lowest rates of *yes* responses to incongruent probes. When there was no relevant prosodic cue in the sentence (neutral prosody) both LC and EC probes were accepted a little over half the time. The *yes/no* judgment data alone cannot tell us exactly how listeners were mentally processing these items; the RT data are also relevant.

The RT data also confirm the role of prosody and the syntactic structure in restoring missing phonemes in the spoken sentence. Incongruency between the prosody of the spoken sentence and the morpho-syntax of the visual probe resulted in slow-downs in RTs to the probe words. Although we had predicted the prosody-congruent probes to be processed faster than prosody-neutral probes, there was no reliable difference between these two conditions in either the lengthened subject or the lengthened VP condition. This could be due to the naturalness of the neutral prosody condition. Pronunciation acceptability judgment data were not collected for the stimuli in the present study because of their acoustic manipulation due to noise replacement. A pronunciation acceptability task for the same sentences with disambiguating phonemes present was reported by D&F (2017), which showed high acceptability ratings for neutral prosody sentences (>98%) comparable to the cooperating prosody sentences (100%). The LC advantage in prosody-neutral probe conditions confirms the effects of syntactic bias in restoring the missing phonemes.

The RT data with respect to the effect of phrase lengths are different from what was observed previously in the 'got it' task of D&F (2017), which showed that both LC and RSH effects influenced the outcomes.

In the present experiment, which employed the same materials as D&F's earlier study (except for the absence of disambiguating phonemes in spoken sentences in the current experiment), results did not fully replicate the earlier findings, specifically in that the current data did not show any RSH effects (see Table 3 for congruent and neutral prosody probes; for incongruent probes see below). There was no advantage for either the LC or the EC probe in the congruent probe conditions. There was an advantage for the LC probe in the prosody-neutral conditions as expected on the basis of the Late Closure strategy. In the prosody-incongruent conditions, for the lengthened subject condition there was an EC advantage, but for the lengthened VP condition there was an LC advantage.

One explanation would be that in the present experiment the role of overt prosodic phrasing in restoring missing phonemes is so compelling that listeners tend to hold on to that interpretation, and are not tempted to revise it based on other less decisive factors such as phrase lengths. (Phrase length effects, e.g., RSH, might occur only when disambiguating morpho-syntax is present in the sentence, perhaps helping the parser to assess how reliable those cues are; see Marslen-Wilson et al. (1992). But that is not relevant to the present concern.)

Let us examine the specific findings in light of this possibility that any potential effects of phrase lengths are outweighed by the influence of prosody. That would explain why in the LC and EC congruent probe conditions there was no advantage for either structure. If the listener had used phrase length patterns as cues to syntactic informativeness of the prosody, an advantage for the unbalanced 4+2 LC structure in the lengthened subject condition, or for the unbalanced 2+4 EC structure in the lengthened VP condition, would be expected, but neither of these was the case.

In contrast, in the neutral prosody condition, there was the usual LC effect but no effect of constituent lengths, although a preference for length balance might have been expected, favoring 3+3 (i.e., EC in the lengthened subject condition and LC in the lengthened VP condition). The fact that this was not so suggests that when there is no overt disambiguating prosody, the main influence on the parser is the Late Closure strategy, rather than a length balance preference. Late Closure would resolve the ambiguity in favor of the LC structure (4+2 in lengthened subject and 3+3 in lengthened VP conditions). This replicates the previous findings in the literature for neutral prosody (e.g., Kjelgaard & Speer, 1999).

Consider now the cases where the probe word conflicts with the analysis based on the prosody of the sentence (and also presumably therefore with the phoneme that was restored online in light of prosody). In this case, when the listener responds *yes*, the parser may be taking the probe word as a prompt, offering an analysis alternative to the one based on prosody on-line, and undertake syntactic reanalysis in order to make the sentence and the probe word compatible. When EC prosody promotes the EC structure online, reanalysis prompted by the conflicting LC probe word would consist of revising to an LC structure. But when LC prosody promotes the LC structure online, the conflicting EC probe word could trigger revising to an EC structure. The former (EC revised to LC) could be expected to be easier than the latter (LC revised to EC) due to the Late Closure strategy. This would explain the LC advantage in the lengthened VP incongruent probe condition.

However, this speculation does not in itself offer an account of why, in the lengthened subject condition, there were faster responses to a conflicting EC probe following a prosodically LC sentence than to an LC probe following a prosodically EC sentence. This unexpected finding needs some explanation. It was not predicted under either RSH, or Late Closure as a syntactic default. We can only suggest that it is due to the fact that the LC prosodic boundary in the prosodically LC sentence was detectable only right before the phoneme-replaced verb, leaving little or no time for the parser to establish its LC analysis before a conflicting probe word appeared. See Stoyneshka et al. (2010) who report a similar effect of recency in their speech shadowing experiment.

Summary of Findings and General Discussion

The present study has shown that Turkish listeners are guided by two of the three factors, syntactic biases and the prosody of the sentence, when restoring noise-replaced phonemes in a late/early closure ambiguity that has not previously been studied from a psycholinguistic perspective. The phoneme restoration paradigm employed in this study confirmed the findings of our previous (2017) study with respect to the role of these two factors, syntax and prosody, in the processing of this ambiguous construction, and also provided new insight into the time-course of these factors. In the phoneme restoration task, the participants were presented with temporarily ambiguous sentences whose disambiguating morphology was masked with noise. The unambiguous probe word with its early closure or late closure morphology was presented very soon (just one word later) after the noise-replaced target word in the sentence, yet response times to acceptance or rejection of it showed sensitivity to the prosodic information in the sentence.

The study also looked for RSH-induced phrase length effects in processing the ambiguity. Although the data indicated immediate use of syntactic and prosodic cues, the on-line results did not show any effects of phrase lengths, as would have been predicted by RSH. This is different from the earlier study of ours, which used the same materials (except for disambiguating morphology being noise-replaced in the present study). That earlier study, employing a 'got it' task, showed an interaction between a syntactic parsing strategy (Late Closure) and phrase lengths (affecting the perceived informativeness of prosodic cues).

The absence of phrase length (RSH) effects in the present study, which employed a measure sensitive to the time-course of parsing decisions, suggests that RSH-induced length effects (as in the prior D&F experiment) may be delayed, as contemplated by Clifton et al.. Clifton and colleagues had observed that in their experiments "One unanswered question involves when listeners make use of prosodic information. If prosodic boundaries are used immediately in processing (e.g., to project a likely syntactic boundary), it may be that the effect of a long phrase following the prosodic boundary (to devalue the interpretive effect of the boundary) is a delayed effect, reflecting a revision that occurs only when the listener is well into the long phrase" (Clifton et al., 2006, pp. 858–859). However, the method employed in their experiments was not able to reveal the time-course of phrase length effects in the perception of informativeness of prosodic cues.

We believe that the method we have employed in the present study tests parsing decisions more rapidly than in the previous experiments on RSH effects. In Clifton et al.'s experiments participants listened to a globally ambiguous sentence (disambiguated by prosody) and then chose one of two sentential paraphrases presented to them visually. This method clearly is offline in its test of parsing decisions. It can also alert participants to the nature of the ambiguity and call upon some explicit cognitive checking process. The less intrusive method employed by D&F was a 'got it' task in which participants simply listened to sentences with cooperating, conflicting and neutral prosody and indicated whether they did or did not understand the spoken sentence. However, in that experiment also, participants were required to provide a response only sentence-finally, after hearing the complete sentence. Thus even that method, though more streamlined than Clifton et al.'s, might also have triggered some post-sentence evaluation process.

By contrast, in the phoneme restoration task proposed by Warren (1970) and employed here, listeners reportedly are unaware that what they were exposed to was not a complete sentence with all phonemes present. Our present data thus imply that RSH effects detected in previous experiments reflected delayed decisions. Most strikingly, in the present study, there was no effect, in the 2+4 length condition, of the early boundary, which would have been expected to lead to an advantage for the EC congruent probe. Further experiments would be needed to explore more precisely the time-course of RSH effects.

Acknowledgements

We are grateful to Martin Chodorow, Luca Campanelli and Cem M. Deniz for their insights on the statistical analyses.

Funding statement

This work was supported by National Science Foundation Dissertation Improvement Grant [grant number 1250473].

References:

Augurzky, P. (2006). *Attaching relative clauses in German: The role of implicit and explicit prosody in sentence processing. MPI Series in Human Cognitive and Brain Sciences*. *MPI series in human cognitive and brain sciences* (Vol. 77). Leipzig.

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

Bashford, J., Riener, K., & Warren, R. (1992). Increasing the intelligibility of speech through multiple phonemic restorations. *Perception & Psychophysics*, *51*(3), 211–217. https://doi.org/10.3758/bf03212247

Bashford, J., & Warren, R. (1987). Multiple phonemic restorations follow the rules for auditory induction. *Attention, Perception, & Psychophysics*, *42*(2), 114–121. https://doi.org/10.3758/bf03210499

Boersma, P., & Weenink, D. (2009). Praat: doing phonetics by computer (Version 5.1.05).

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

Deniz, N. D., & Fodor, J. D. (2017). Phrase Lengths and the Perceived Informativeness of Prosodic Cues in Turkish. *Language and Speech*, *60*(4), 505–529. https://doi.org/10.1177/0023830916665653

Eckstein, K., & Friederici, A. D. (2006). It's early: Event-related potential evidence for initial interaction of syntax and prosody in speech comprehension. *Journal Of Cognitive Neuroscience*, *18*(10), 1696–1711. https://doi.org/10.1162/jocn.2006.18.10.1696

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*. (Unpublished doctoral dissertation). University of Connecticut, Storrs.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Colthart (Ed.), *Attention and performance XII: The psychology of reading*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

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

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, *14*(2), 178–210. https://doi.org/10.1016/0010-0285(82)90008-1

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

Göksel, A., & Kerslake, C. (2005). *Turkish: A comprehensive grammar*. New York: Routledge.

Güneş, G. (2013). *Limits of Prosody in Turkish*. *Dilbilim Arastirmalari Dergisi (The Journal of Linguistic Research)* (Vol. 2013/1).

İşsever, S. (2003). Information structure in Turkish: The word order-prosody interface. *Lingua*, *113*(11), 1025–1053. https://doi.org/10.1016/s0024-3841(03)00012-3

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

Kan, S. (2009). *Prosodic domains and syntax prosody mapping in Turkish*. (Unpublished M.A. thesis). Boğaziçi University, Istanbul.

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. https://doi.org/10.1006/jmla.1998.2620

Kornfilt, J. (2011). The sentential subject constraint/CED as a left-dislocation constraint in Turkish. *Proceedings of the 7th Workshop on Altaic Formal Linguistics (WAFL 7)*, 203–219.

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

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods*, *49*(4), 1494–1502. https://doi.org/10.3758/s13428-016-0809-y

Magen, H. S. (1997). The extent of vowel-to-vowel coarticulation in English. *Journal of Phonetics*, *25*(2), 187–205. https://doi.org/http://dx.doi.org/10.1006/jpho.1996.0041

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. https://doi.org/10.1007/bf01708576

Nash, R. (1973). *Turkish intonation*. Paris: Mouton, the Hague.

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. https://doi.org/10.1162/jocn.2011.21610

Powers, G. L., & Wilcox, J. C. (1977). Intelligibility of temporally interrupted speech with and without intervening noise. *The Journal of the Acoustical Society of America*, *61*(1), 195–199. https://doi.org/doi:http://dx.doi.org/10.1121/1.381255

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

R Core Team. (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL 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.

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.

Selkirk, E. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge: MIT Press.

Selkirk, E. (2011). The syntax-phonology interface. In J. Goldsmith, J. Riggle, & A. Yu (Eds.), *The Handbook of phonological theory*. Oxford: Blackwell.

Snedeker, J., & Trueswell, J. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language*, *48*(1), 103–130. https://doi.org/http://dx.doi.org/10.1016/S0749-596X(02)00519-3

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. https://doi.org/10.1007/bf01708573

Steinhauer, K. (2003). Electrophysiological correlates of prosody and punctuation. *Brain and Language*, *86*(1), 142–164. https://doi.org/10.1016/S0093-934X(02)00542-4

Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain responses indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience*, *2*, 191–196.

Steinhauer, K., & Friederici, A. D. (2001). Prosodic boundaries, comma rules, and brain responses: The closure positive shift in ERPs as a universal marker for prosodic phrasing in listeners and readers. *Journal of Psycholinguistic Research*, *30*(3), 267–295. https://doi.org/10.1023/a:1010443001646

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.

Truckenbrodt, H. (1995). *Phonological phrases: Their relation to syntax, focus and prominence.* (Unpublished doctoral dissertation). Masachusetts Institute of Technology, Cambridge, MA.

Tunley, A. (1999). *Coarticulatory influences of liquids on vowels in English*. (Unpublished doctoral dissertation). King's College, University of Cambridge, Cambridge.

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

Warren, P., Grabe, E., & Nolan, F. (1995). Prosody, phonology and parsing in closure ambiguities. *Language and Cognitive Processes*, *10*(5), 457–486. https://doi.org/10.1080/01690969508407112

Warren, R. (1970). Perceptual restoration of missing speech sounds. *Science*, *167*(3917), 392–393.

Warren, R., & Sherman, G. (1974). Phonemic restorations based on subsequent context. *Attention, Perception, & Psychophysics*, *16*(1), 150–156. https://doi.org/10.3758/bf03203268

Watt, S. M., & Murray, W. S. (1996). Prosodic form and parsing commitments. *Journal of*

Psycholinguistic Research, *25*(2), 291–318. https://doi.org/10.1007/bf01708575

Webman-Shafran, R., & Fodor, J. D. (2016). Phrase length and prosody in on-line ambiguity resolution. *Journal of Psycholinguistic Research*, *45*, 447–474. https://doi.org/10.1007/s10936-015-9358-2