The Effect of Prosodic Representation on Agreement and Reading Comprehension in L1 and L2 Speakers

Elizabeth Pratt

1. Introduction

This study examines the effect of manipulating implicit prosody projection in light of several L2 processing models, comparing L1 and late L2 learners of English in comprehension and grammatical processing measures during reading.

1.1. L2 processing models

Studies of second language (L2) learning have long observed that in contrast to L1 acquisition, ultimate attainment in an L2 is highly variable. One prevalent area of discussion considers whether L1 and L2 processing are fundamentally different, and if so, where these differences may arise. Below, I summarize three prominent models of L2 processing, each of which attributes nonconvergence with L1 speakers to different sources: representation of the grammar (Shallow Structure Hypothesis), processing mechanisms (Declarative-Procedural Model), and processing capacity (‘Good Enough’ Hypothesis).

1.1.1. Shallow Structure Hypothesis

Clahsen and Felser's (2006) Shallow Structure Hypothesis (SSH) claims that L2 processing is qualitatively different from native processing in that L2 speakers rely more on lexical and semantic information rather than complete computation of syntactic information. In this model, native speakers have two processing routes at their disposal—a full parsing route and a shallow processing route. The full parsing route allows computation of complex syntactic phrase structure while the shallow route does not. According to Clahsen and Felser, basic sentence comprehension only requires segmentation of the input and semantic integration of the components. This can often be accomplished using semantic and pragmatic knowledge, without recourse to complex syntactic information. However, a more detailed phrase structure representation must be activated to process purely structural relations, or to make use of elements such as intermediate traces.

While in theory L2 speakers may have access to both routes, this requires the development of a native-like grammatical representation, which is constrained by critical period effects in acquisition. Thus, to the degree that L2 learners can attain a native-like grammatical representation, they can also make use of the full processing route. However, for most adult learners, critical period effects would not allow this, and so they are dependent on semantic or plausibility information to compensate their deficient syntactic representations.

While the SSH is able to account for findings that L2 speakers have difficulty integrating complex syntactic information, and do not transfer L1 processing strategies to the L2 (Keating, 2009; Marinis et al., 2005; Papadopoulou & Clahsen, 2003; Sato & Felser, 2006), these observed differences do not necessarily equate evidence for a qualitatively different representation. Citing wh-extraction data from Marinis et al. (2005), Clahsen and Felser propose that while natives construct a complex representation of the incoming sentence complete with intermediate traces, L2 speakers may process incoming material much more simply, relying on semantic/pragmatic information to assign thematic roles as soon as
possible. The resulting structure does not project any intermediate traces, which may account for erroneous thematic assignment and other processing errors (Juffs, 2005; Williams, Möbius, & Kim, 2001).

Clahsen and Felser acknowledge these performance effects, despite evidence of L2 speakers’ knowledge of subjacency constraints. However, if L2 speakers do possess this grammatical knowledge, or utilize an alternate yet equally complex grammatical construal (e.g., pro-binding, as discussed in Dekydtspotter & Sprouse, 2003), it is difficult to argue that they can never deploy this knowledge during parsing, or that their grammatical representation is necessarily shallow or syntactically simple.

### 1.1.2. Declarative-Procedural Model

An alternative to variation in grammatical representation are models which point to differences in how that knowledge is accessed during processing. Models such as the Declarative-Procedural (DP) model (Ullman, 2004) draw upon two types of memory systems, the declarative and the procedural. In native linguistic processing, the two systems roughly align with the mental lexicon and the mental grammar: the declarative system handles lexical storage—form and meaning, irregular morphological forms, as well as other memorized forms. The procedural system handles sequential or hierarchical linguistic structures, drawing on syntax, regular morphological forms, and other implicit knowledge.

Due to maturational factors, late L2 learners may rely more heavily on declarative knowledge, using formulae and memorized forms to compensate for their procedural learning deficits. While this allows them to develop an extensive explicit knowledge of an L2’s vocabulary and grammar, it predicts difficulty in constructions that are not easily memorizable, such as long-distance dependencies. Despite this declarative-procedural imbalance, Ullman claims that with sufficient exposure and proficiency, L2 learners may eventually converge with natives, automatizing grammatical processes in procedural memory.

The DP model is notable in that it draws upon well-established neurophysiological memory models. However, it is unclear on how learners transfer from declarative to procedural knowledge, and why implicit methods of learning can input directly into procedural memory for some learners but not others (Brill-Schuetz & Morgan-Short, 2014; Morgan-Short et al., 2014).

### 1.1.3. ‘Good-Enough’ Hypothesis

A third type of model is the ‘Good Enough’ (GE) hypothesis, which draws on variations in processing load and capacity (Ferreira, Bailey, & Ferraro, 2002; Ferreira, 2003). The GE is premised on the assumption that most of the time, listeners are not required to perform a complete parse of the input, and so tend to generate representations based on both syntactic information and general heuristics. These heuristics prioritize thematic templates, pragmatics, animacy, and plausibility over morphological and syntactic processing. This strategy may be responsible for misinterpretations based on integrating elements in the input before all relevant structural information is available. Christianson and colleagues (Christianson et al., 2001; Christianson, Zacks, & Ferreira, 2006) cite evidence from comprehension studies of garden-path sentences, indicating that initial misparses persist, despite ultimately correct reanalyses.

Beyond this effect, comprehenders may also misinterpret noncanonical sentences if the interpretation clashes with plausibility factors and/or an expectation of agent-verb-patient structure. Ferreira and colleagues (Ferreira & Patson, 2007; Ferreira, 2003) argue that the real-time constraints on processing—including time constraints, working memory limitations, or other processing load factors—result in GE processing patterns. Comprehenders may sacrifice complete syntactic parsing, applying pragmatic or thematic template expectations, to economize these resources.

In terms of L2 processing, the GE model would predict greater and earlier reliance on heuristics over full parsing. L2 processing is typically slower and more error-prone than native processing, suggesting that it is more cognitively demanding. Whether due to working memory limitations, overall proficiency, or reduced automaticity in the L2, the greater processing load would manifest in heavier reliance on plausibility and templates.
Lim (2011) compared the predictions of the SSH and the GE models in L2 processing, looking at performance during comprehension and translations tasks. While the L2 speakers were more susceptible than native speakers to errors based on plausibility, they were able to compute both syntactic and semantic information in a pattern similar to native speakers, suggesting perhaps quantitative, but not qualitative differences. Task goals also affected performance: when processing for translation, L2 speakers were more attentive to structure than when processing for comprehension alone. This argues directly against the predictions of the SSH, which claims that grammatical representations are either accessible to the L2 speaker or are not.

Crucially, the GE does not propose a different grammatical representation or architecture for native and L2 speakers. While the SSH and DP models suggest that there are fundamental differences in how language is represented and processed in an L2 (at least after the critical period), the GE claims that for both groups the fuller syntactic algorithm is available along with semantic-pragmatic heuristics, but processing strategy or depth is dependent on the processing goals and task load.

While not directly incorporated into the GE model, it is plausible that working memory and other processing constraints interact with the selection of parsing strategies. The additional demand of processing under noise conditions or within an L2 may promote greater reliance on heuristics in an attempt to conserve resources.

2. Prosody and processing
2.1. Prosody in silent reading

Within the general psycholinguistics literature, both oral prosody and implicit prosody (the prosody projected onto text during silent reading), have been shown to influence parsing decisions, and case memory and task load during processing.

Explicit prosody has been shown to make use of phrasing (breaks) and intonational cues (pitch accents) (Schafer et al., 1996) to disambiguate between syntactic representations. In the absence of explicit prosodic cues, is there evidence that phonological features such as phrasing and intonation are projected and can similarly influence interpretation? Evidence drawing from the Baddeley and Hitch (1974) model, and Slowiaczek and Clifton (1980) suggests that if subvocalization is blocked, comprehension during silent reading is impaired; thus, it appears that this form of rehearsal involves projection of phonological/prosodic information, and that this information contributes to the processing and comprehension.

In the years following, there has been additional evidence that fluent readers are not only able to produce prosody during oral reading, but while reading silently as well (Bader, 1998; Fodor, 2002). The experience of hearing an ‘inner voice’ during silent reading has long been anecdotally attested, however, recent research supports this experience theoretically and empirically, suggesting a number of functions which it may perform.

Koriat et al. (2002) propose the Structural Precedence Hypothesis, which claims that during reading, readers establish an early structural frame for a phrase or sentence based on function words and morphosyntactic cues which may indicate general phrase structure. Prosody during oral reading may reflect this early processing and be used to help maintain a structure in memory during integration (Kjelgaard & Speer, 1999). This prosodic representation may then provide the initial framework for syntactic and semantic parsing (Schafer, 1997; Speer, Shih, & Slowiaczek, 1989).

Other researchers have suggested that prosody more directly influences both early syntactic and reanalysis processes. Bader (1998) shows that prosody is able to affect the ease of reanalysis during reading of ambiguous sentences, claiming that during reading, both a prosodic and a syntactic structure are produced. If revision of syntactic structure is necessary, it is made more difficult if the prosodic structure must be revised as well (Prosodic Constraint on Reanalysis, Bader 1998: 8). The Implicit Prosody Hypothesis takes this claim further by stating that a default prosodic contour is projected onto text during silent reading, and this projection directly affects ambiguity resolution (Fodor, 1998, 2002).

Similar claims have been made for lexical stress variations (Breen & Clifton, 2005), rhythmic stress patterns (Ashby & Clifton, 2005; Ashby & Martin, 2008; Kentner, 2012), and prosodic phrase lengths (Hirose, 2003; Hwang & Schafer, 2009; Hwang & Steinhauer, 2011), suggesting that multiple aspects of prosody may influence processing during reading.
2.1.1. Implicit prosody and agreement

Disruption of implicit prosody has also been shown to affect agreement processing, where errors in subject-verb agreement are less likely to be detected if natural reading rhythm is impeded. Kreiner (2005) proposed that the processing difficulty resulting from a subject-verb mismatch reflects online syntactic integration: natural reading prosody facilitates this integration, and allows subject-verb mismatches to be more easily detected. However, if this integration is disrupted in some way, mismatches will be more difficult to detect, and thus less processing difficulty will be seen.

Testing this hypothesis, Kreiner compared eye movements of participants reading sentences using natural prosody, with those using a three-word fixed grouped prosody which did not align with the phrasal structure of the experimental sentences.

Additionally, to test the prediction that natural prosody helps to maintain a structural representation in memory during integration, Kreiner contrasted an ‘adjacent’ condition (1a), where the subject and verb were immediately adjacent, with a ‘distant’ condition (1b), where the subject and verb were separated by a relative clause:

(1)  
   a. Adjacent match/mismatch:  
      The audience who have been waiting excitedly to this fashion show watch how the model/s step proudly on the stage.  
   b. Distant match/mismatch:  
      The audience watch with a smile how the model/s dressed in a ridiculous hat step proudly on the stage.

There was no significant effect of prosody when the subject and verb were adjacent; in the distant condition however, a mismatch effect was found for the natural prosody, but not when grouped prosody was applied. These results suggest that natural prosody does facilitate agreement processing, particularly when the processing and/or working memory load is greater.

2.1.2. Prosodic phrasing and comprehension

Pedagogical research has investigated how manipulation of text presentation may enhance reading skill and support the development of reading fluency. Pre-segmentation of text into meaningful phrasal chunks has been shown to improve the reading performance of both children (LeVasseur et al., 2006; O’Shea & Sindelar, 1983) and less skilled adult readers (Cromer, 1970). Skilled readers seem to be more resistant to text segmentation, suggesting that their own phrasing skills override cues from the input.

In an early study, Cromer (1970) investigated the effect of text segmentation on the reading comprehension of advanced, intermediate, and poor readers. Sentences were presented in one of four formats: whole sentence (2a), single word (2b), phrase (2c), or fragmented group (2d).

(2)  
   a. The cow jumped over the moon.  
   b. The | cow | jumped | over | the | moon.  
   c. The cow jumped | over the moon.  
   d. The cow | jumped over the | moon.

Assuming that skilled readers chunk text into phrases while less skilled readers read word-by-word, Cromer predicted that guiding the less skilled groups to read in phrases would improve their comprehension (i.e., make them look like skilled readers), and guiding the skilled readers to read word-by-word would disrupt their comprehension (i.e., make them look like less skilled readers).

Cromer found that comprehension patterns varied based on both skill level and presentation format. Advanced readers were unaffected by presentation format, and comprehended equally well in all conditions. Intermediate readers were disrupted in the word and fragment conditions, but improved with

---

1 English translations of the Hebrew sample sentences, as given in Kreiner (2005).
the phrase condition. Poor readers comprehended best in the word format and were not significantly affected by any of the other formats.

Cromer concluded that while advanced and poor readers may be impervious to text manipulations, either due to the strength of their own phrasing skills (advanced readers), or the deficit in their reading fluency (poor readers), intermediate-level readers may benefit from textual phrasing as an aid to reading fluency. This suggests that as long as sufficient vocabulary skills have been acquired, facilitating text presentation may directly influence reading comprehension.

In terms of major prosodic features, languages perhaps most notably differ in intonation and lexical stress patterns (see discussion in Cutler 2012), but prosodic phrasing patterns are more universal, at least when those patterns align with syntactic constituents. In several notable studies, even with no previous exposure to the test language, listeners were able to correctly identify pauses occurring at constituent boundaries (Endress & Hauser, 2010; Pilon, 1981; Wakefield, Doughtie, & Yom, 1974) In fact, L2 speakers may rely more heavily on prosodic rather than syntactic cues, particularly when syntax and prosody are misaligned (Harley, Howard, & Hart, 1995). Evidence would suggest, then, that where the performance of L1 and L2 speakers may diverge is not necessarily in prosodic phrasing itself, but in its relation to overall fluency and the availability of processing resources. A task manipulating the prosodic projection environment would thus be further complicated in reading, adding the processing burden of word recognition and integration in a second language.

### 3. Methods

#### 3.1. Participants

For this study, 63 native English speakers (mean 21.6 yrs; 37 female, 26 male) and 24 Spanish-English late bilinguals (L2) were recruited (mean 22.8 yrs; 14 female, 10 male). All were Queens College students enrolled in an introductory psychology course and received course credit for their participation. The final pool of L2 participants was selected based on age of acquisition and arrival information gathered from a background questionnaire.

#### 3.2. Materials

Sixty-four experimental items were distributed across 4 lists, each also containing 16 practice items and 128 fillers. Experimental materials consisted of relative clause sentence sets in which an intervening plural attractor noun either matched or mismatched the target verb. The nouns and target verb were selected from a list of the 5000 most frequent words of each type in the Corpus of Contemporary American English (COCA), and all head and attractor nouns were animate. The 2x2 design crossed the factors of structural complexity (Simple, Complex) and grammaticality (Grammatical, Ungrammatical). Sentences within each set were matched for length using either an adjunct modifier for the simple sentences, or an additional relative clause for the complex sentences (see Table 1). Grammaticality was manipulated by varying the number feature of the main verb. A sample set of experimental materials appears in Table 1.

<table>
<thead>
<tr>
<th>Complexity</th>
<th>Grammaticality</th>
<th>Sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple</td>
<td>Grammatical</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>Ungrammatical</td>
<td>b</td>
</tr>
<tr>
<td>Complex</td>
<td>Grammatical</td>
<td>c</td>
</tr>
<tr>
<td></td>
<td>Ungrammatical</td>
<td>d</td>
</tr>
</tbody>
</table>

Table 1: Sample set of materials
3.3. Procedure

Participants were tested in a single session, and all items were presented electronically using E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA). Each participant was randomly assigned to a presentation paradigm and stimulus list upon recruitment.

In the SENTENCE presentation, sentences were presented individually on one line in their entirety, and reading was self-paced. This format most closely imitates natural reading, providing a baseline of comparison for the other presentation formats. Each sentence was preceded by a fixation cross which appeared centrally on the screen for 1000 ms. The sentence then appeared and remained on the screen until the participant pressed the space bar.

In the WORD presentation, sentences were presented word-by-word, in rapid serial visual presentation (RSVP) format, at a fixed rate of 500 ms per word. While RSVP has often been utilized as a method of enhancing reading speed (Rahman & Muter, 1999), presenting sentences at a rate of two words per second is strikingly slow, and has been shown to disrupt processing during reading (Fernández, 2007). This effect has been attributed to the fixed presentation rate of the materials, suggesting that the invariable pace interferes with the projection of prosody onto the text (Castelhano & Muter, 2001; Fernández, 2007).

In the PHRASE presentation, sentences were presented in three phrasal segments, with a break after the head noun and again after the relative clause (see (3)), and reading was self-paced. The first segment was preceded by a fixation cross which appeared centrally on the screen for 1000 ms. The first segment of the sentence then appeared and remained on the screen until the participant pressed the space bar to advance to the next segment.

(3)   a. Simple: The reporter | who called the senators every once in a while | write(s)…
     b. Complex: The reporter | who called the senators that Scott supported | write(s)…

For all presentation paradigms, following each sentence, participants were prompted to rate the sentence on a 6-point Likert scale (1 = “very bad”, 6 = “perfect”). To minimize low ratings due to the complexity of the sentences, participants were instructed to rate the sentences based on whether a 300-level English professor would consider them grammatical. Participants were then prompted to respond to a true/false comprehension probe, and received speed and accuracy feedback on their responses. Participants were instructed to respond as quickly and accurately as possible, and were allowed 5000 ms to rate each sentence, and an additional 5000 ms to respond to the comprehension probe.

Following the main experimental session, participants completed a language background questionnaire, as well as reading proficiency tests in English, and in both English and Spanish for the L2 speakers.

In consideration of the GE, we predicted that overall processing load would be greater for L2 speakers than for L1 speakers. Also, assuming that (proficient) readers project prosodic phrasing onto text, disrupting or facilitating this function may affect processing load, with consequential effects on comprehension and grammatical processing.

4. Results

For the L1 speakers, there was a significant effect of presentation type on comprehension (F(2,60) = 5.96, p<.004), with pairwise comparisons indicating that comprehension errors were greater in the WORD condition than the SENTENCE condition (t(40) = 2.58, p<.02, see Fig 1). In the grammaticality judgments, there was also a significant effect of presentation type (F(2,60) = 6.20, p<.04), with paired comparisons indicating that the difference between ratings for grammatical versus ungrammatical sentences was significant in the PHRASE condition (Simple: t(20) = 3.87, p<.0001; Embedded: t(20) = 3.61, p<.001), but not in the SENTENCE or WORD conditions (see Figs 2 and 3).
For the L2 speakers, there was a marginally significant effect of presentation type on comprehension ($F(2,21) = 2.73, p<.08$), with numerical trends indicating that comprehension errors were lower in the WORD condition than in the SENTENCE condition (see Fig 4). There was no significant effect of presentation type on agreement error detection, indicating that the L2 speakers were not consistently able to detect errors in any of the three formats (see Figs 5 and 6).
5. Discussion

This study examined the effect of manipulating implicit prosody on comprehension and grammatical processing during reading. Results showed that text presentation mediates processing load during reading differently in L1 and L2 populations, but these differences are likely due to reading...
proficiency and parsing strategies. High proficiency (L1) readers may tend to project appropriate prosodic phrasing and contours onto text during reading, making a phrase-based presentation facilitative to processing and a slow word-by-word presentation disruptive. For lower proficiency (L2) speakers, a word-by-word presentation showed evidence of being facilitative to comprehension, while there was no effect of presentation format on grammatical processing. These findings allow us to make several broad inferences regarding L2 processing and the role of prosody in processing.

Drawing from the GE hypothesis, parsing strategies of both L1 and L2 speakers will prioritize comprehension over complete syntactic analysis. However, there is evidence that a base level of lexical access must be achieved before prosody may be incorporated into the parsing strategy, and before presentation format can influence it to affect both comprehension and grammatical processing (LaBerge & Samuels, 1974; Perfetti, 1985, 1988; Rasinski, 1999; Schwanenflugel et al., 2004). Thus, in the comprehension measures, the L1 speakers performed relatively well in the baseline sentence condition, and so were not further facilitated by the phrase presentation. The word-by-word presentation however, disrupted processing for the L1 group but showed evidence of facilitating the L2 group. This format may force more resources to the initial step of lexical access, allowing lower proficiency readers to better integrate incoming materials.

For more proficient readers (e.g., L1 speakers), grammatical processing is typically not prioritized when taskload is high, but as resources were freed by reducing load (here, due to phrase presentation), L1 grammatical processing improved. For lower proficiency readers (e.g., L2 speakers), processing capacity was still taxed regardless of presentation, and so no effect was found.

This study not only supports a strong role of prosody in both comprehension and grammatical processing during reading, but also points to critical differences between L1 and L2 processing during reading. These differences are not necessarily qualitative, but relate to proficiency level and the distribution of resources based on processing load. Further inquiries will correlate these results with measures of oral prosody, and further clarify the relationship between prosodic projection, proficiency, and processing.

References


