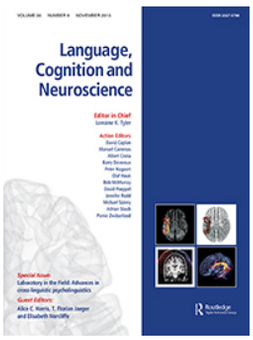

Routledge
Taylor & Francis Group

Language, Cognition and Neuroscience

ISSN: 2327-3798 (Print) 2327-3801 (Online) Journal homepage: http://www.tandfonline.com/loi/plcp21

Contextual predictability and the prosodic realisation of focus: a cross-linguistic comparison

Rory Turnbull, Rachel Steindel Burdin, Cynthia G. Clopper & Judith Tonhauser

To cite this article: Rory Turnbull, Rachel Steindel Burdin, Cynthia G. Clopper & Judith Tonhauser (2015) Contextual predictability and the prosodic realisation of focus: a cross-linguistic comparison, Language, Cognition and Neuroscience, 30:9, 1061-1076, DOI: 10.1080/23273798.2015.1071856

To link to this article: http://dx.doi.org/10.1080/23273798.2015.1071856

Language, Cognition and Neuroscience, 2015
Vol. 30, No. 9, 1061–1076, http://dx.doi.org/10.1080/23273798.2015.1071856

Routledge
Taylor & Francis Group

Contextual predictability and the prosodic realisation of focus: a cross-linguistic comparison

Rory Turnbull, Rachel Steindel Burdin, Cynthia G. Clopper* and Judith Tonhauser

Department of Linguistics, Ohio State University, 1712 Neil Ave, 222 Oxley Hall, Columbus, OH 43210, USA

(*Received 1 July 2013; accepted 6 July 2015*)

This study explored the effect of contextual predictability on the prosodic realisation of focussed expressions in American English and Paraguayan Guaraní. Pairs of native speakers played an interactive game to elicit utterances that varied in the location of focus in the NP and whether this location was predictable from visual context. The English results confirmed that focussed expressions had more rising pitch accents, longer durations, and higher f0 than non-focussed expressions. Differences between focussed and non-focussed expressions were enhanced when the location of focus was not predictable from context. The Guaraní results confirmed that focussed expressions had distinctive pitch accent and duration patterns relative to non-focussed expressions. Overall prosodic prominence was enhanced when the location of focus was not predictable from context. These results, which are discussed within information-based theories of language production, suggest contextual predictability affects the prosodic realisation of focus, and that this predictability-dependence varies across languages.

Keywords: prosody; focus; contextual predictability; Paraguayan Guaraní; American English

The realisation of linguistic material is affected by linguistic and non-linguistic context in many subdomains, ranging from small phonetic detail (Baese-Berk & Goldrick, 2009), to morphological encoding (Gundel, Hedberg, & Zacharski, 1993; Leonetti, 2004), to the prosodic marking of syntax (Snedeker & Trueswell, 2003; cf. Speer, Warren, & Schafer, 2011), among others. Under the assumption that context affects the realisation of linguistic elements at all levels of linguistic structure, the first objective of the current study was to examine the effect of context in a new linguistic domain, namely the prosodic realisation of focus. Specifically, we sought to determine whether the predictability of the location of focus in a NP influences the prosodic realisation of that focus. The second objective was to explore cross-linguistic differences and similarities in the prosodic realisation of focus and in the relationship between contextual predictability and prosodic form. The languages of investigation were American English and Paraguayan Guaraní (Tupí-Guaraní), two genetically distinct languages. In each language, experimental participants directed a confederate to place a series of tiles depicting different objects into numbered boxes on a game board. The available items, which were visible to both interlocutors, were manipulated to vary the degree to which the visual context provided cues for the listeners to the location of focus in the critical NP of the talkers' utterances. In this way, the location of the focus was, for the confederate listeners, either predictable or unpredictable from the context.

Previous cross-linguistic work on the relationship between contextual predictability and linguistic form

Research on the relationship between contextual predictability and linguistic form has demonstrated that contextual predictability is correlated with phonetic reduction. Specifically, more predictable elements are more likely to be reduced than less predictable elements (Aylett & Turk, 2004; Bell, Brenier, Gregory, Girand, & Jurafsky, 2009; Bell, Jurafsky, Fosler-Lussier, Girand, & Gregory, 2003; Clopper & Pierrehumbert, 2008; Raymond, Dautricourt, & Hume, 2006; Van Son & Pols, 2003). Similarly, lexical frequency has been found to correlate with phonetic reduction in many languages, including English (Gahl, 2008), Taiwan Southern Min (Myers & Li, 2009), European Portuguese (Vigário, 2003), and Dutch, Finnish, and Russian (Van Son, Bolotova, Lennes, & Pols, 2004). Additionally, Catalan word-final devoicing is fully neutralising when the intended meaning is predictable from context, but only partial when the intended meaning is not predictable (Charles-Luce, 1993). That is, an acoustic difference between word-final obstruents is maintained in ambiguous contexts, but the distinction is abandoned once sufficient context for distinguishing meaning is provided. Likewise, vowel epenthesis in Dutch is less likely to occur in contextually more predictable words than in less predictable words (Tily & Kuperman, 2012).

Morphological and lexical elements are similarly affected by contextual predictability. Gundel et al. (1993) demonstrated that in English, Japanese, Mandarin, Russian, and Spanish, the more salient a referent is in

*Corresponding author. Email: clopper.1@osu.edu

discourse, the shorter and less complex the referring expression is likely to be. These effects are further mediated by higher-level narrative context, in both English (Fowler, Levy, & Brown, 1997) and Thai (Vajrabhaya & Kapatsinski, 2011). In addition, case-marking on grammatical objects is more likely for objects that are less contextually salient, an effect which is putatively universal (Aissen, 2003). Across these studies, the same generalisation arises: increased contextual predictability leads to smaller linguistic forms.

This generalisation has been interpreted in terms of information-based theories of language production, such as Aylett and Turk's (2004) smooth signal redundancy hypothesis, Levy and Jaeger's (2007) and Jaeger's (2010) uniform information density hypothesis, and Pluymaekers, Ernestus, and Baayen's (2005) informational redundancy hypothesis. These hypotheses posit that talkers manipulate linguistic information in response to the contextual predictability of their utterances, with the aim of creating a relatively constant or "smooth" amount of information transmission over time. Essentially, these hypotheses view speech production as an optimisation problem where the key variables are talker effort and listener understanding, and predict an inverse relationship between linguistic form and contextual predictability. That is, when more linguistically relevant contextual information is available to the listener, less linguistic information is provided by the talker.

If, as hypothesised, contextual predictability affects the realisation of linguistic elements at every level of linguistic structure, predictability effects ought to be observable in the prosodic domain. Most previous work on the role of context in the realisation of prosodic prominence has focussed on the realisation of different types of information structure (see, e.g. Braun, 2006; Breen, Fedorenko, Wagner, & Gibson, 2010; Krahmer & Swerts, 2001; Swerts, Krahmer, & Avesani, 2002 on Germanic and Romance languages). A notable exception to this work on prosody and information structure is Watson, Arnold, and Tanenhaus' (2008) study of the role of contextual predictability in the acoustic realisation of prosodic prominence in American English. In a game of tic-tac-toe (noughts and crosses), participants announced their moves to each other. Highly predictable moves, such as blocking an opponent's winning move, were generally produced with shorter duration and smaller f0 excursion than less predictable moves.

This previous work on contextual predictability, in both the prosodic and segmental domains, motivates our specific hypothesis that across languages increased contextual predictability about which expression is focussed will lead to weaker prosodic marking of focus. More specifically, our hypothesis concerns cases where the focus of an utterance is either contextually predicted or not contextually predicted. We expect fewer or weaker cues to the focussed expression, or a smaller difference between focussed and non-focussed expressions, when the focus is contextually predictable. That is, if the visual context establishes that the noun of a NP will be focussed, our hypothesis is that the noun will be realised with fewer or weaker cues to focus than if the visual context does not predict that the noun will be focussed, i.e. if noun focus is not contextually predictable. Thus, unlike much of the previous work described above, we are not directly concerned with the predictability of the linguistic elements themselves (e.g. how predictable the adjective or the noun is in the NP), but rather how predictable the location of the focus is on an expression in the NP.[1] Since the relationships between contextual predictability and linguistic form are purported to be universal, such effects should be observable across languages. However, the way in which these effects manifest themselves may be language-specific, due to cross-linguistic differences in, for instance, prosodic typology, word order, and marking of information structure. Further, as the prosodic realisation of focus has been found to be a variable process, we expect that some aspects of this realisation may be influenced by changes in contextual predictability, but that the effect of contextual predictability may not be observed for all prosodic correlates of focus within or across languages. The experiment reported in this paper tests this hypothesis for two languages, American English and Paraguayan Guaraní.

Why American English and Paraguayan Guaraní?

English has been the primary language for research on both the prosodic realisation of focus and predictability effects on linguistic form, and as such serves as an excellent point of comparison for cross-linguistic research. Paraguayan Guaraní (henceforth Guaraní) was selected as the comparison language because its overall prosodic structure is similar to English, thus allowing for a comparative analysis. However, as described below, Guaraní differs from English with respect to the prosodic realisation of focus, in terms of both the phonological pitch accent contrasts and the phonetic factors employed. Thus, together, these two languages provide a useful test case for exploring potential cross-linguistic differences in the effect of contextual predictability on the prosodic realisation of focus.

Background on the prosodic systems of American English and Paraguayan Guaraní

English stress is lexically specified, and stressed syllables can receive pitch accents. The prosodic system of American English features a relatively large pitch accent inventory (Jun, 2014) and two levels of phrasing, including

intermediate and intonational phrases that are marked on their right edge by phrase accents and boundary tones, respectively (Beckman, Hirschberg, & Shattuck-Hufnagel, 2005; Ladd, 2008; Pierrehumbert & Hirschberg, 1990). Deaccenting of given material is common in American English (Bolinger, 1972; Ladd, 1980), and has been argued to be an obligatory process (Selkirk, 1995). However, the empirical data on deaccenting rates in American English suggest that it is a variable process, affected by factors such as word class, overall accentability of the word, and underlying metrical structure (Calhoun, 2010; Ito & Speer, 2006; Katz & Selkirk, 2011; see also Féry & Kügler, 2008; Riester & Piontek, in press on German).

Paraguayan Guaraní is a Tupí-Guaraní language spoken in Paraguay, where it is an official language alongside Spanish. The language is agglutinative and mildly polysynthetic (Gregores & Suárez, 1967; Velázquez-Castillo, 2004). In Guaraní, as in English, stress is lexically specified (Gregores & Suárez, 1967), and stressed syllables can receive pitch accents (Clopper & Tonhauser, 2013). Clopper and Tonhauser (2013) described Paraguayan Guaraní prosody within the autosegmental-metrical framework, and identified rising (LH) and falling (HL) pitch accents, and high (H%) and low (L%) boundary tones which occur at the right edges of intonational phrases. Although Clopper and Tonhauser (2013) did not find evidence of deaccenting in two-word utterances, Burdin et al. (in press), in an analysis of a subset of the data presented in the current study, observed deaccenting within NPs at rates similar to or exceeding deaccenting in American English NPs. Neither Clopper and Tonhauser (2013) nor Burdin et al. (in press) found evidence for a level of phrasing below the intonational phrase. Thus, unlike English, Guaraní exhibits a relatively small pitch accent inventory and only a single level of prosodic phrasing above the level of the word. Guaraní is a relatively understudied language and the descriptions presented by Clopper and Tonhauser (2013) and Burdin et al. (in press) constitute the only literature on Guaraní intonational phonology. As such, the results of the current study help to augment our understanding of the language, in addition to testing hypotheses about the relationship between contextual predictability and linguistic form.

Our specific hypothesis concerns the effects of contextual predictability of the location of focus in a NP on the prosodic realisation of the focussed expression. We assume that a "focus" is an information-structurally prominent expression. Specifically, an expression is a focus if it answers the explicit or implicit question under discussion (Roberts, 2012). Thus, we use the term "focus" to indicate a pragmatic property of expressions, rather than a phonetic or phonological property. In English, focus may be realised with an expanded pitch range (Eady & Cooper, 1986) or with a L + H* pitch accent (Pierrehumbert & Hirschberg, 1990). Additionally, focussed expressions are lengthened relative to non-focussed expressions (Eady & Cooper, 1986). Our hypothesis is that increased contextual predictability about which expression is focussed will lead to weaker prosodic marking of focus. Given this hypothesis, we expected to observe weaker prosodic marking of focus when it is predictable from the visual context which expression is focussed compared to when it is not predictable from the visual context which expression is focussed. Specifically, in English, we expected to observe fewer rising pitch accents on focussed expressions, shorter duration of focussed expressions, and lower f0 peaks associated with focussed expressions when the expression that is focussed is predictable from the visual context compared to when the location of focus is not predictable from the visual context. In Guaraní, focussed expressions are more likely to be realised with a rising (LH) than a falling (HL) pitch accent, and those with a rising pitch accent may be realised with a more steeply rising f0 contour. Focussed expressions are also lengthened relative to non-focussed expressions, as in English (Clopper & Tonhauser, 2013). Therefore, in Guaraní we expected to observe more rising pitch accents, longer duration, and steeper f0 slopes associated with focussed expressions when the expression that is focussed is not predictable from the visual context compared to when the location of focus is predictable from the visual context.

Methods

An interactive task was used with English- and Guaraní-speaking participants in the USA and Paraguay, respectively. The task required a "Director" (a native speaker participant) to instruct a "Follower" (a native speaker confederate) to place a series of tiles depicting differently coloured items into numbered boxes on game boards. The available items and the order of items to be placed in the boxes were manipulated to create a fully crossed design examining the prosodic realisation of focussed adjectives, nouns, and NPs in visual contexts from which the location of focus was either predictable or not predictable.

Participants

American participants were recruited from Ohio State University and received partial course credit for their time. Data were analysed from 10 participants (6 female), who were all monolingual native speakers of English and ranged in age from 18 to 21 years old ($M = 19$). One female speaker (age 27) served as the Follower.

Guaraní-speaking participants were recruited from San Lorenzo, Paraguay, and were paid for their time. Data were analysed from 10 participants (5 female), who were all bilingual speakers of Guaraní and Spanish, with self-reported native proficiency in Guaraní. The participants

ranged in age from 18 to 51 years old ($M = 33$). Two female speakers (ages 49 and 51) served as the Followers.

Stimulus materials

The stimulus materials consisted of game boards with four or six numbered boxes and game tiles. The Director received boards with coloured items in each of the boxes, as shown in Figure 1(a) for a game board with four numbered boxes. The Follower received boards with blank numbered boxes, as shown in Figure 1(c). The game tiles each depicted one of the coloured items, as shown in Figure 1(b). In English, the items were a flower, a lion, a deer, an owl, and a train, and each item was orange, green, blue, brown, or yellow, for a total of 25 unique tiles. In Guaraní, the items were a dog (*jagua*), a pig (*kure*), a bird (*guyra*), a tree (*yvyra máta*), and a flower (*yvoty*), and each item was white (*morotĩ*), black (*hũ*), red (*pytã*), yellow (*sayju*), or blue (*hovy*), for a total of 25 unique tiles.

Procedure

The Director and the Follower were seated at a table facing each other. On each trial, the Director and the Follower were given filled and blank boards, respectively. The Follower could see the Director's face, but not the Director's board. A subset of the 25 tiles was placed on the table between them and the Director was asked to tell the Follower which tiles to put into which boxes. The Directors were instructed to always use the same frame sentence that included both the colour adjective and the shape noun. Sample English and Guaraní sentences are given in (1) and (2), respectively; note that the adjective follows the noun in Guaraní. The Directors were not explicitly instructed to fill the boxes in order, but all Directors did so. Upon hearing the Director's utterance, the Follower selected the appropriate tile from the table and placed it in the appropriate box. Once the board was complete, the Director confirmed its correctness. We refer to the event of filling a board as a trial; each trial consisted of four or six utterances.

(1) Put the orange flower in box two.

(2)

E-moi	kure	hũ	peteĩ-me.
you-put	pig	black	one-in

"Put the black pig in box one."

The available subset of tiles on each trial was manipulated to create two visual context conditions: which expression (the noun, adjective, or NP) was focussed in all of the utterances in the trial was either predictable or not predictable from the context. In the Predictable Context, each board had four boxes, and five tiles were placed between the Director and the Follower. The five tiles were selected such that the expression of the NP that was going to be focussed in the Director's utterances was predictable from the selected tiles. In Predictable Context Adjective Focus trials, exemplified in Figure 1, the five tiles had the same shape but differed in colour (e.g. orange flower, green flower, blue flower, etc.), i.e. it was predictable that the adjectives were going to be focussed in the Director's utterances. In Predictable Context Noun Focus trials, the five tiles had the same colour but differed in shape (e.g. orange flower, orange lion, orange deer, etc.), i.e. it was predictable that the nouns were going to be focussed in the Director's utterances. In Predictable Context NP Focus trials, the five tiles differed from one another in both shape and colour (e.g. orange flower, green lion, blue deer, etc.), i.e. it was predictable that the entire NPs

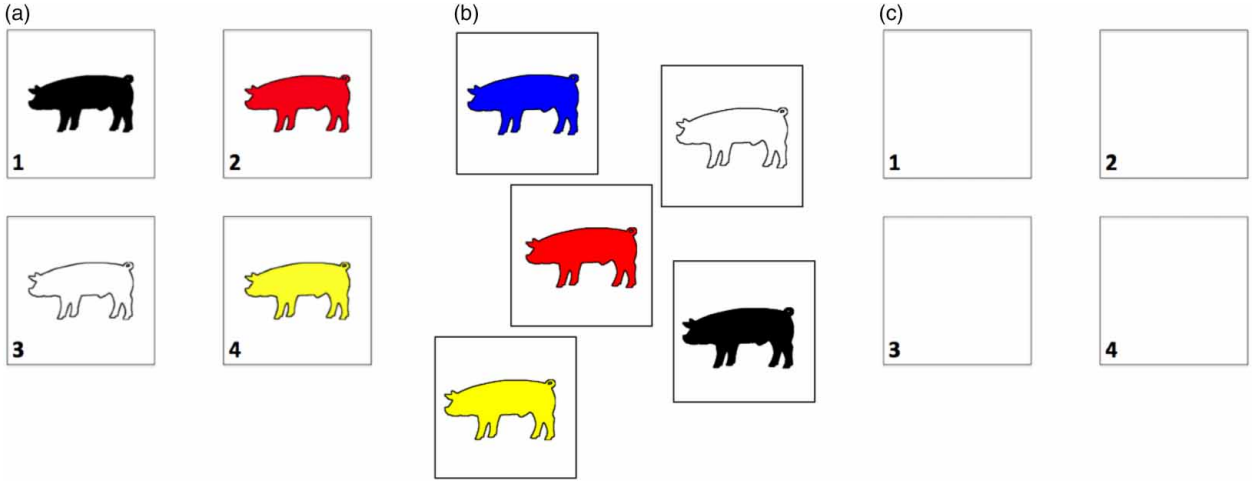

Figure 1. Sample Guaraní Director's game board (a), game tiles (b), and Follower's game board (c). The Director's game board contains four pigs, coloured black, red, white, and yellow. The game tiles include five pigs, coloured blue, white, red, black, and yellow. [To view this figure in colour, please see the online version of this Journal.]

were going to be focussed in the Director's utterances. Thus, in the Predictable Context trials, which expression in the Director's utterances was focussed was fully predictable from the visual context (i.e. the tiles placed between the experiment participants). For a more detailed description of the Predictable Context trials in this experiment, see Burdin et al. (in press).

In the Unpredictable Context, each board had six boxes and eight tiles were placed between the Director and the Follower. These eight tiles included the six items on the Director's board and two items that were selected to control the colour and shape similarity of available tiles throughout the trial. Some tiles were unique in both shape and colour, and others shared either shape or colour with other available tiles. Therefore, in the Unpredictable Context trials, whether the adjective, the noun, or the entire NP would be focussed in the Director's utterances was not predictable from the remaining tiles.

Thus, in contrast to the Predictable Context trials, in which both the Director and the Follower knew from the visual context which part of the NP would be focussed in the Director's utterance, in the Unpredictable Context trials, only the Director had this information prior to producing his or her utterance. For example, if the tile to be placed (T_x) differed from the tile that had just been placed (T_{x-1}) in colour alone, then in the utterance for tile T_x, the adjective was focussed because the adjective of the utterance for tile T_x contrasted with the adjective of the utterance for tile T_{x-1} (e.g. *green owl … [brown]$_F$ owl,* where [ ]$_F$ marks the focussed expression). Similarly, if tile T_x differed from tile T_{x-1} in shape alone, then the noun was focussed in the utterance for tile T_x, because it contrasted with the noun of the utterance for tile T_{x-1} (e.g. *green owl … green [deer]$_F$*). Finally, if tile T_x differed in both shape and colour from tile T_{x-1}, then the entire NP was focussed in the utterance for tile T_x, because it contrasted with the NP of the utterance for tile T_{x-1} (e.g. *green owl … [brown deer]$_F$*). Crucially, the Follower did not know which part of the NP was focussed prior to hearing the Director's utterance for any given tile in this condition. Thus, in the Unpredictable Context trials, information about which part of the NP was focussed was only available to the Follower from the Director's utterance, whereas in the Predictable Context trials, this information was also available from the visual context. We therefore expected to see differences in the prosodic realisation of focus in the two context conditions, reflecting this difference in the availability of information about focus location in the visual context across conditions.

In Unpredictable Context trials, utterances in all three focus conditions were realised on each trial. Each set of six boxes included two fillers, and the first box of a trial was always a filler. Of the remaining four boxes, two were NP Focus targets, one was an Adjective Focus target, and one was a Noun Focus target.[2] As in the Predictable Context, the NP Focus targets in the Unpredictable Context were unique in colour and shape among the available tiles and the Noun and Adjective Focus targets always had at least one competitor tile with the same colour or shape, respectively, among the remaining tiles.[3]

Each session began with a block of four Predictable Context NP Focus trials, followed by four trials in each of the Predictable Context Adjective Focus and Predictable Context Noun Focus blocks. The order of these two blocks was counterbalanced across participants. NP Focus trials were completed first so that all five distinct colours and all five distinct shapes were presented to each participant on the first trial. Finally, participants completed eight trials in the Unpredictable Context. Within each block, the order of the trials was randomised separately for each participant.

The Directors wore a head-mounted microphone and their utterances were digitally recorded. Prior to analysis, utterances that were produced with errors or disfluencies were removed (93 for English; 96 for Guaraní). The analysis is therefore based on 707 English utterances and 704 Guaraní utterances.

Analysis

Each target utterance was segmented for analysis into an individual sound file, identified by block, trial number, and box number. The block labels were opaque to focus condition (i.e. A, B, C, and D). The English utterances were transcribed using the Tones and Break Indices (ToBI) conventions for Mainstream American English (Beckman & Ayers Elam, 1997), which distinguish different pitch accent heights (e.g. L* vs. !H* vs. H*), different pitch accent alignments (e.g. L* + H vs. L + H*), and two levels of prosodic phrasing above the word (e.g. L- vs. L%). The Guaraní utterances were transcribed in the autosegmental-metrical framework, following the analysis proposed by Clopper and Tonhauser (2013), as described in the introductory section on the prosodic system of Guaraní above. Each utterance in each language was transcribed by one author and checked by another author to ensure reliability (ToBI transcriptions for American English are generally reliable; Pitrelli, Beckman, & Hirschberg, 1994; Syrdal & McGory, 2000). From these transcriptions, the pitch accents of the adjectives and nouns were extracted. The presence or absence of a phrase boundary (i.e. intonational phrase boundaries in Guaraní and both intermediate and intonational phrase boundaries in English) was also coded for two positions: within the NP (between the noun and adjective) and following the NP. Adjective and noun word boundaries were identified by visual and auditory inspection of the spectrogram and waveform, and word durations were extracted automatically from the boundary placement. For the English data, where the

pitch accents were overwhelmingly high or rising, the f0 peak associated with the high target of pitch-accented nouns and adjectives was extracted to provide a phonetic measure of f0 height. This f0 peak did not always occur within the word; late peaks were identified consistent with the phonological transcription. For the Guaraní data, in which falling and rising pitch accents were observed, the absolute value of the slope of the f0 rise or fall associated with each pitch accent was extracted. The absolute f0 slope measure was selected because Clopper and Tonhauser (2013) found it to be a better predictor of focus than f0 range (i.e. absolute excursion size). Slope was calculated as the absolute difference in Hz between the f0 maximum and the f0 minimum, divided by the temporal distance between them in seconds. Slope can thus be interpreted as rate of change in Hz/s. All f0 peaks and valleys were identified by hand to avoid segmental effects; the measures were extracted via an automatic script, with every measurement hand-checked to correct for errors arising from non-modal voice quality.

The phonological and phonetic data were analysed using mixed-effects regression models predicting accenting, accent type, duration, and f0 peak (for English) or f0 slope (for Guaraní). Separate models were constructed for the adjectives and nouns. Each model had focus condition (Adjective, Noun, NP), context condition (Predictable, Unpredictable), and their interaction as fixed effects. The fixed effects were treatment-coded with Noun Focus and Adjective Focus as the reference levels for focus condition for the noun and adjective analyses, respectively. The Predictable Context was the reference level for context condition in all analyses. Random intercepts were included for Director, noun, and adjective. To avoid model over-specification, the maximal random slopes that resulted in convergence and did not include strong random effect correlations ($|r| > .9$) were retained for each model (Barr, Levy, Scheepers, & Tily, 2013). The analyses of pitch accent type and f0 were restricted to the subset of accented tokens. The English pitch accents were collapsed into two categories for analysis – high (H*, !H*) and rising (L + H*, L* + H) – to allow for stronger comparison with Guaraní, which only exhibits two pitch accent categories. To control for phonological prosodic effects on the phonetic variables, pitch accenting (accented or unaccented) and following phrase boundary (present or absent) were included as fixed effects in the word duration analyses, and pitch accent type (high or rising in English, rising or falling in Guaraní) was included as a fixed effect in the f0 analyses.[4] To control for the contribution of duration to the f0 slope measure, duration was included as a fixed effect in the f0 slope analysis. Statistical significance of planned pairwise comparisons was determined assuming a normal distribution of model z and t estimates; absolute z and t values greater than 1.96 were interpreted as significant ($p < .05$).

Results

Based on previous research on focus marking in English and Guaraní, we expected to observe main effects of focus condition for each dependent variable (e.g. nouns are longer in Noun Focus than Adjective Focus). Additionally, given our specific hypothesis regarding the effect of contextual predictability on prosodic prominence, an interaction between focus and context condition was predicted, such that focussed items in the Unpredictable Context were expected to have stronger or more prominent cues to focus than focussed items in the Predictable Context. Although English and Guaraní mark focus through somewhat different phonetic and phonological means, both languages were expected to exhibit the same general pattern of focus and predictability effects.

American English

The percentages of rising (L + H*, L* + H), high (H*, !H*), and unaccented adjectives and nouns in the English utterances are shown in the top and bottom panels of Figure 2, respectively. The results of the mixed-effects models for English adjectives and nouns are summarised in Tables 1 and 2, respectively. Adjectives were more likely to be unaccented in Noun Focus than Adjective Focus, consistent with variable deaccenting of non-focussed expressions in English (Calhoun, 2010; Ito & Speer, 2006; Katz & Selkirk, 2011). Among the accented adjectives, rising pitch accents were marginally less frequent in Noun and NP Focus than Adjective Focus in the Predictable Context, but rising pitch accents were significantly less frequent in Noun Focus than Adjective Focus in the Unpredictable Context. This pattern is consistent with the use of rising pitch accents to mark focussed expressions in English (Pierrehumbert & Hirschberg, 1990) and suggests that the difference in pitch accent types across focus conditions was larger in the Unpredictable Context than the Predictable Context, as predicted.

Similarly, nouns were more likely to be unaccented in Adjective Focus than Noun Focus, consistent with variable deaccenting of non-focussed expressions in English (Calhoun, 2010; Ito & Speer, 2006; Katz & Selkirk, 2011). Among accented nouns, rising pitch accents were more frequent in Noun Focus than Adjective or NP Focus, as expected. However, rising pitch accents on the noun were less frequent in Noun Focus in the Unpredictable Context than the Predictable Context, and the difference between Noun Focus and Adjective Focus was smaller in the Unpredictable Context than the Predictable Context. Thus, unlike for adjectives, the difference in pitch accent types across focus conditions was greater in the Predictable Context than the Unpredictable Context, contrary to expectation.

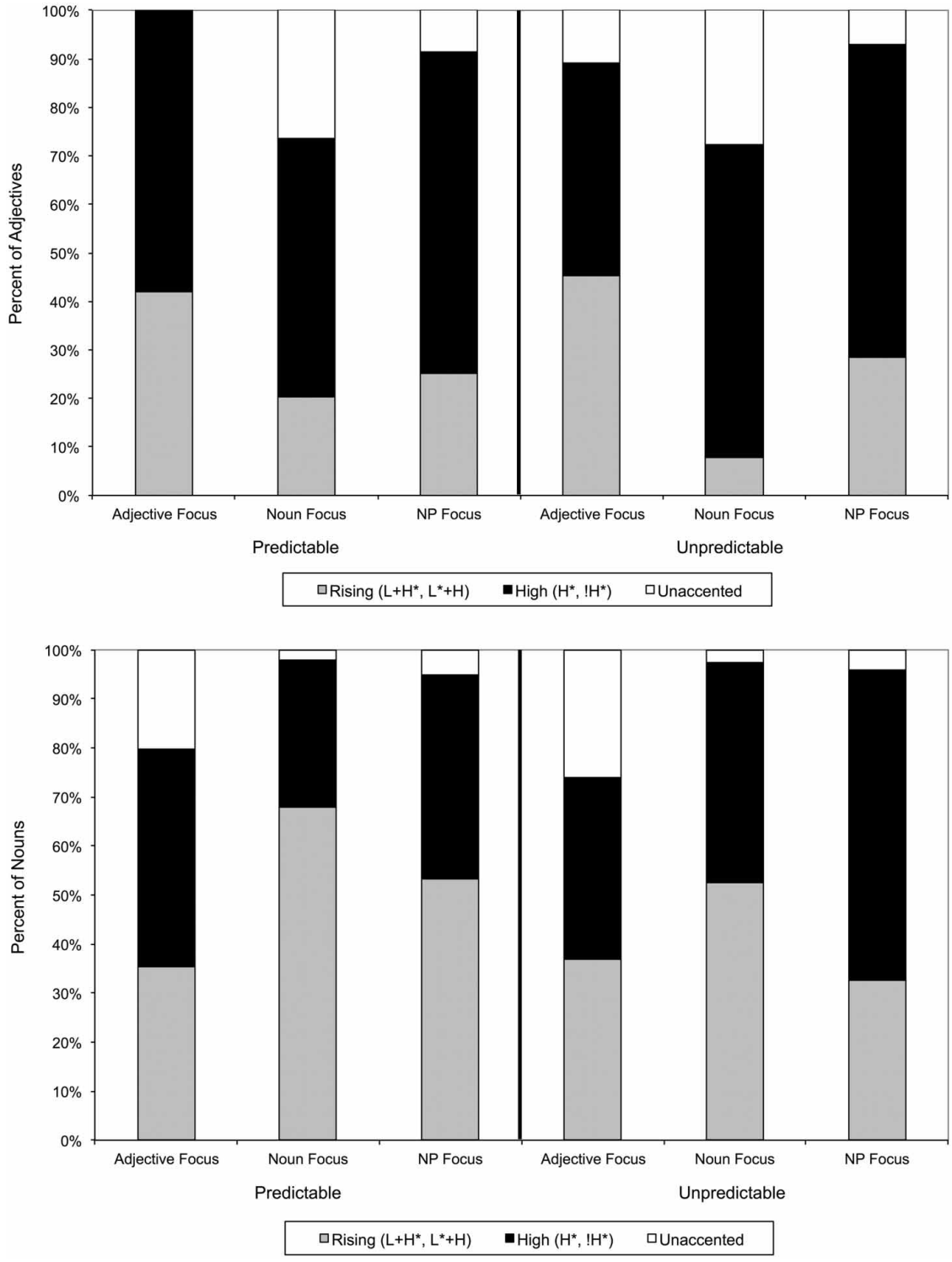

Figure 2. Percentages of rising (L + H*, L* + H), high (H*, !H*), and unaccented adjectives (top) and nouns (bottom) in English for each focus condition in each context condition.

Focus and context condition also affected noun duration in English. As shown in the right panel of Figure 3, nouns were longer in Noun Focus than Adjective Focus, consistent with lengthening of focussed expressions in English (Eady & Cooper, 1986). Nouns were also longer in the Unpredictable Context than the Predictable Context, but this context effect was reduced in NP Focus, leading to a greater duration difference between Noun

Table 1. Summary of the context and focus effects in the mixed-effects models for English adjectives.

	Accenting	Accent type (rising)	Duration	f0 peak
Context (Unpred.)	$\beta = 16.65$, $z = 0.02$	$\beta = 0.75$, $z = 1.58$	$\beta = 0.66$, $t = 0.06$	**$\beta = 9.91$, $t = 2.49$**
Focus (Noun)	**$\beta = -1.67$, $z = -2.64$**	$\beta = -1.33$, $z = -1.71$	$\beta = 3.32$. $t = 0.43$	$\beta = 0.13$, $t = 0.03$
Focus (NP)	$\beta = 0.03$, $z = 0.05$	$\beta = -1.34$, $z = -1.85$	$\beta = 9.02$, $t = 1.23$	$\beta = 2.54$, $t = 0.61$
Unpred. × Noun	$\beta = -16.54$, $z = -0.02$	**$\beta = -3.33$, $z = -3.56$**	$\beta = -10.77$, $t = -0.98$	**$\beta = -14.45$, $t = -2.55$**
Unpred. × NP	$\beta = -16.39$, $z = -0.02$	$\beta = -0.36$, $z = -0.64$	$\beta = -11.94$, $t = -1.20$	**$\beta = -13.83$, $t = -2.11$**

Notes: Significant effects are in bold. "Unpred." refers to the Unpredictable Context, and the × symbol indicates interactions between factors.

Table 2. Summary of the context and focus effects in the mixed-effects models for English nouns.

	Accenting	Accent type (rising)	Duration	f0 peak
Context (Unpred.)	$\beta = -0.15$, $z = -0.15$	**$\beta = -0.82$, $z = -2.42$**	**$\beta = 25.00$, $t = 3.02$**	**$\beta = 7.21$, $t = 2.54$**
Focus (Adjective)	**$\beta = -2.57$, $z = -3.90$**	**$\beta = -1.44$, $z = -3.39$**	**$\beta = -37.79$, $t = -5.18$**	$\beta = -5.93$, $t = -1.47$
Focus (NP)	$\beta = -0.94$, $z = -1.28$	**$\beta = -0.69$, $z = -2.05$**	$\beta = 0.92$, $t = 0.13$	$\beta = -0.18$, $t = -0.05$
Unpred. × Adjective	$\beta = -0.14$, $z = -0.13$	**$\beta = 1.15$, $z = 2.24$**	$\beta = 14.14$, $t = 1.20$	**$\beta = -9.55$, $t = -2.20$**
Unpred. × NP	$\beta = 0.47$, $z = 0.42$	$\beta = -0.35$, $z = -0.79$	**$\beta = -22.02$, $t = -2.05$**	**$\beta = -15.38$, $t = -4.15$**

Notes: Significant effects are in bold. "Unpred." refers to the Unpredictable Context, and the × symbol indicates interactions between factors.

Focus and NP Focus in the Unpredictable Context than the Predictable Context. This pattern reflects greater duration differences across focus conditions in the Unpredictable Context than the Predictable Context, as expected. As shown in the left panel of Figure 3, no effects of focus or context condition were observed for adjective duration. The effects of focus and context condition on noun duration were independent of phonological pitch accenting and phrasing. Both adjectives and nouns were longer when they were accented than when they were unaccented (adjectives: $\beta = 22.52$, $t = 3.40$; nouns: $\beta = 18.40$, $t = 2.23$) and when they were followed by a phrase break than when they were not (adjectives: $\beta = 99.76$, $t = 8.95$; nouns: $\beta = 36.46$, $t = 5.88$).

Focus and context condition also affected f0 peak height for both adjectives and nouns in English, as

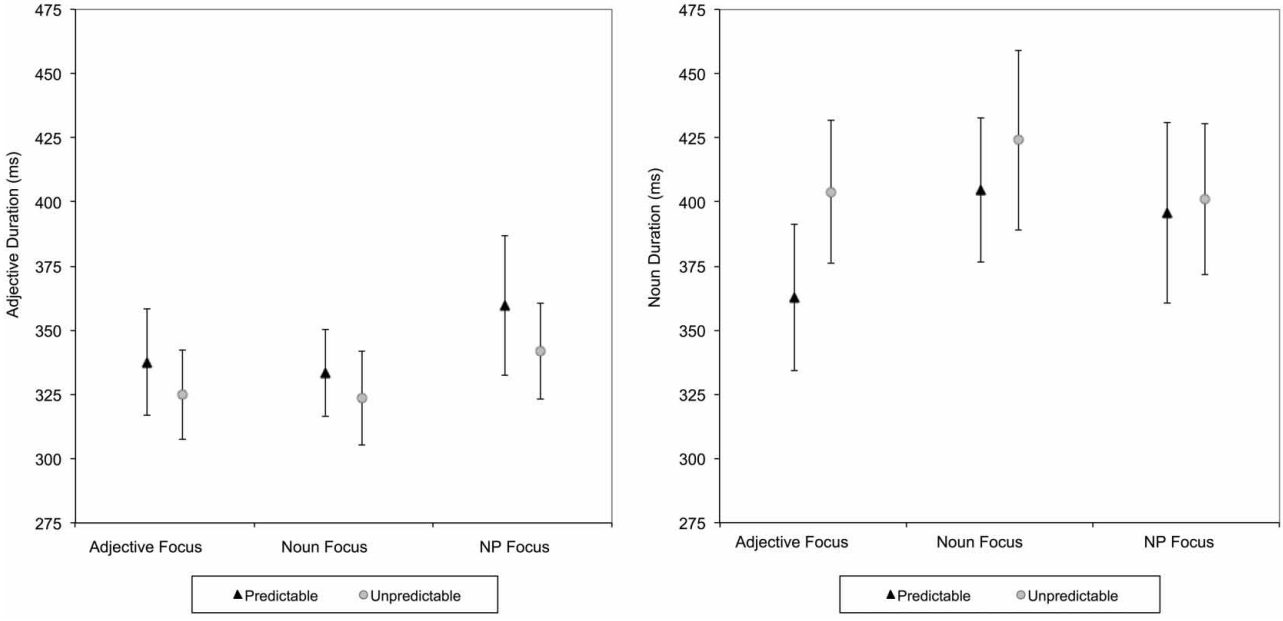

Figure 3. Mean adjective (left) and noun (right) duration in English for each focus condition in each context condition. Error bars are standard error of Director means.

shown in the left and right panels of Figure 4, respectively. For accented adjectives, the f0 peak was higher in the Unpredictable Context than the Predictable Context, but this context effect was reduced in Noun and NP Focus relative to Adjective Focus. Similarly, for accented nouns, the f0 peak was higher in the Unpredictable Context than the Predictable Context, but this context effect was reduced in Adjective and NP Focus relative to Noun Focus. This interaction between focus and context condition reflects higher f0 peaks associated with focussed expressions than non-focussed expressions (Eady & Cooper, 1986), with a larger focus effect in the Unpredictable Context than the Predictable Context, as expected. For both nouns and adjectives, these effects of focus and context condition on f0 peak were independent of overall effects of pitch accent type: f0 peaks were higher for rising pitch accents than for high accents for both nouns and adjectives (adjectives: $\beta = 16.88$, $t = 9.16$; nouns: $\beta = 15.67$, $t = 8.66$).

Taken together, these results demonstrate the expected effects of focus and contextual predictability on pitch accenting, duration, and f0 peak in English. Focussed expressions were more likely to be realised with rising pitch accents, whereas non-focussed expressions were more likely to be unaccented. Focussed nouns were longer than non-focussed nouns and focussed expressions were realised with higher f0 peaks than non-focussed expressions. The effects of focus on adjective pitch accent type, noun duration, and adjective and noun f0 peak were larger in the Unpredictable Context than the Predictable Context, as expected. The effect of focus on noun pitch accent type was smaller in the Unpredictable Context than the Predictable Context, contrary to our predictions. Context condition had no significant effect on deaccenting. Thus, our hypothesis that the prosodic cues to focus would be enhanced in the Unpredictable Context relative to the Predictable Context was confirmed for some of our measures for some focussed expressions. We also observed one effect in the opposite direction to our prediction. However, the lack of an effect of context condition on adjective duration and noun or adjective deaccenting across focus conditions is not strong evidence against our hypothesis because we expected these effects to vary both within and across languages. We have therefore obtained evidence for this variability within English and consider cross-linguistic variation in our discussion of the Guaraní results below.

Paraguayan Guaraní

The percentages of rising (LH), falling (HL), and unaccented nouns and adjectives in the Guaraní utterances are shown in the top and bottom panels of Figure 5, respectively. The results of the mixed-effects models for Guaraní nouns and adjectives are summarised in Tables 3 and 4, respectively. Nouns were more likely to be unaccented in Adjective Focus than Noun Focus, suggesting deaccenting of non-focussed expressions. Among the accented nouns, fewer rising pitch accents were observed in NP Focus than Noun Focus, consistent with Clopper and Tonhauser's (2013) observation of more rising pitch accents on focussed expressions than non-focussed expressions.

Adjectives were rarely unaccented and no effects of focus or context condition were observed on adjective

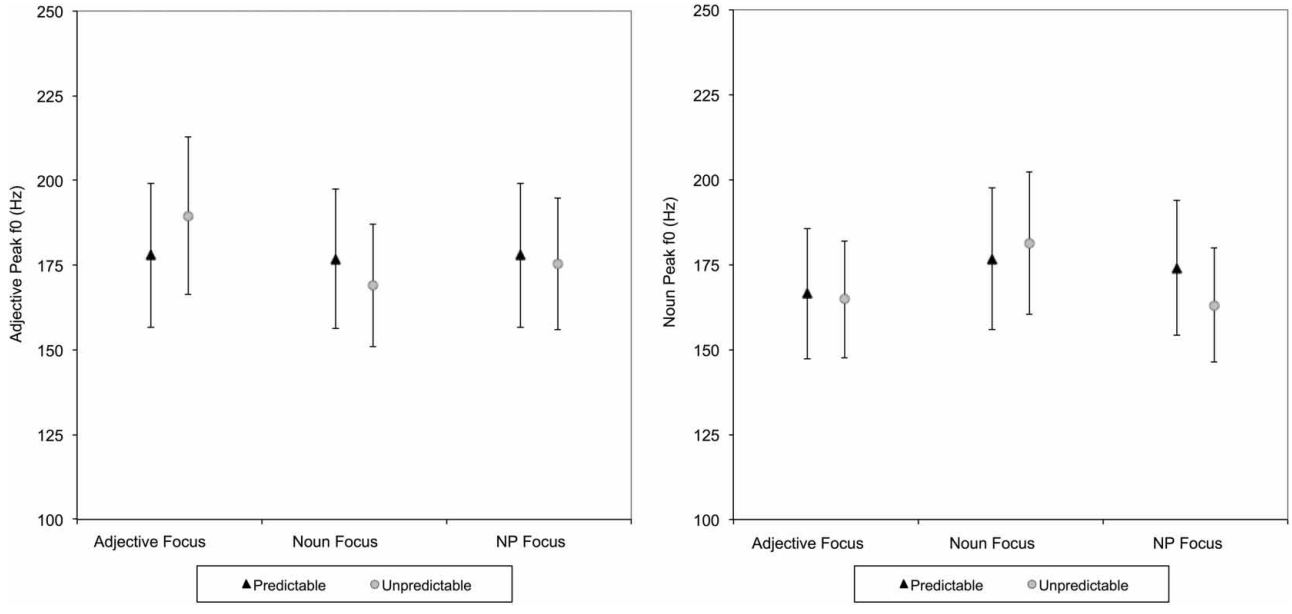

Figure 4. Mean f0 peak for accented adjectives (left) and nouns (right) in English for each focus condition in each context condition. Error bars are standard error of Director means.

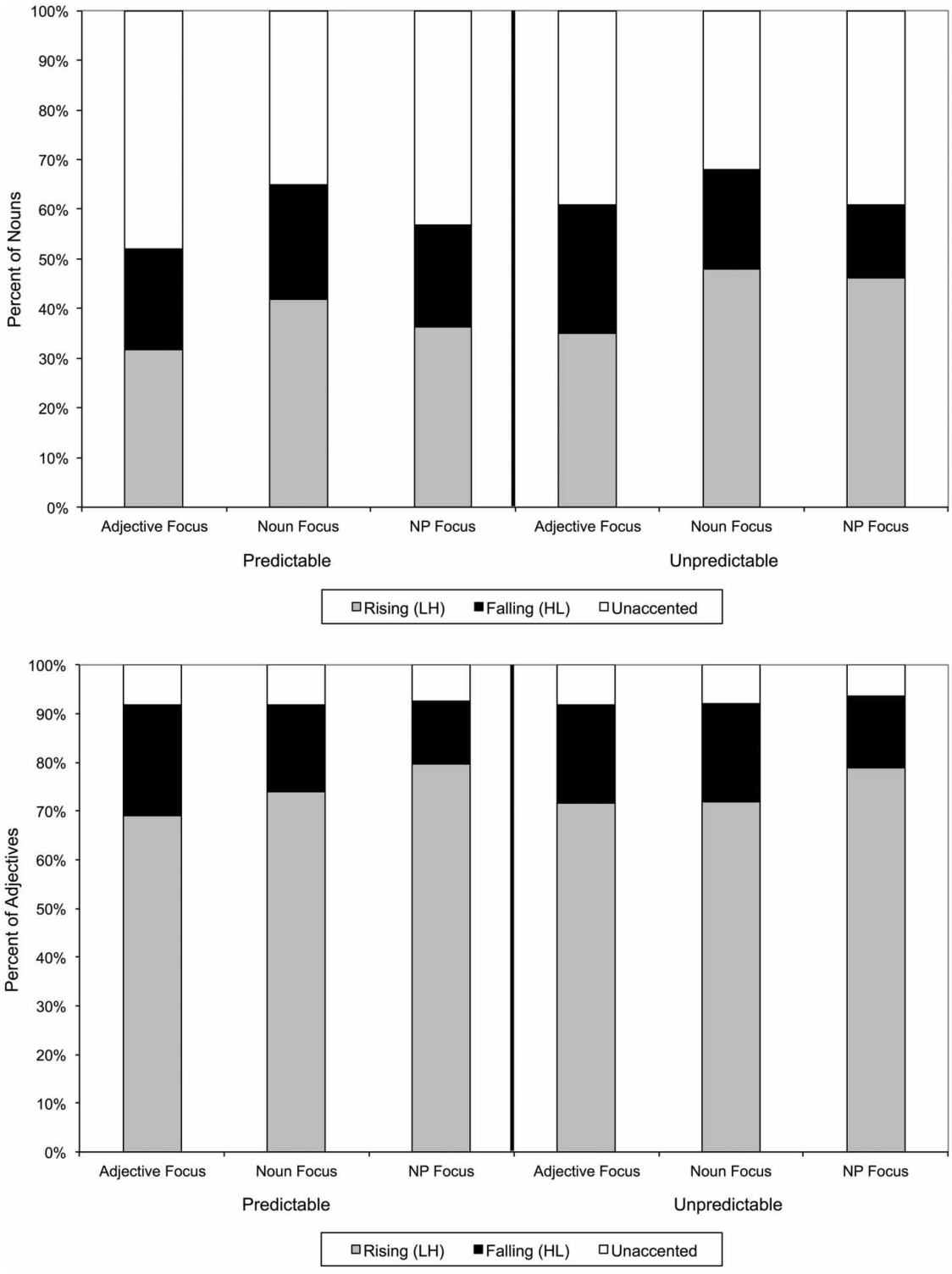

Figure 5. Percentages of rising (LH), falling (HL), and unaccented nouns (top) and adjectives (bottom) in Guaraní for each focus condition in each context condition.

accenting. Among the accented adjectives, more rising pitch accents were observed in Noun and NP Focus than Adjective Focus. Thus, unlike the pattern observed for nouns, focussed adjectives were realised with falling pitch accents more often than non-focussed adjectives. Figure 5 shows that adjectives were realised with rising pitch accents more often overall than nouns (74% of adjectives vs. 39% of nouns). The falling pitch accent may

Table 3. Summary of the context and focus effects in the mixed-effects models for Guaraní nouns.

	Accenting	Accent type (LH)	Duration	f0 slope
Context (Unpred.)	$\beta = 0.46$, $z = 0.77$	$\beta = -0.26$, $z = -0.35$	$\beta = 3.48$, $t = 0.24$	$\beta = -11.21$, $t = -0.86$
Focus (Adjective)	**$\beta = -1.19$, $z = -1.99$**	$\beta = -0.40$, $z = -0.81$	**$\beta = -30.60$, $t = -2.51$**	$\beta = -16.77$, $t = -1.39$
Focus (NP)	$\beta = -0.29$, $z = -0.62$	**$\beta = -1.03$, $z = -2.00$**	$\beta = 0.45$, $t = 0.05$	$\beta = -9.74$, $t = -0.82$
Unpred. × Adjective	$\beta = 0.51$, $z = 0.76$	$\beta = -0.24$, $z = -0.30$	$\beta = 8.85$, $t = 0.67$	$\beta = 13.11$, $t = 0.68$
Unpred. × NP	$\beta = -0.63$, $z = -1.02$	$\beta = 1.16$, $z = 1.54$	$\beta = -1.36$, $t = -0.11$	$\beta = 15.00$, $t = 0.85$

Notes: Significant effects are in bold. "Unpred." refers to the Unpredictable Context, and the × symbol indicates interactions between factors.

Table 4. Summary of the context and focus effects in the mixed-effects models for Guaraní adjectives.

	Accenting	Accent type (LH)	Duration	f0 slope
Context (Unpred.)	$\beta = 0.46$, $z = 0.71$	$\beta = 1.46$, $z = 1.20$	$\beta = 8.32$, $t = 0.71$	**$\beta = 46.72$, $t = 2.89$**
Focus (Noun)	$\beta = 0.23$, z = 0.43	**$\beta = 0.81$, $z = 1.98$**	$\beta = 13.63$, $t = 1.72$	$\beta = 13.00$, $t = 0.95$
Focus (NP)	$\beta = 0.08$, $z = 0.15$	**$\beta = 0.94$, $z = 2.17$**	**$\beta = 23.63$, $t = 2.80$**	$\beta = -7.45$, $t = -0.55$
Unpred. × Noun	$\beta = -0.30$, $z = -0.33$	$\beta = 0.25$, $z = 0.31$	$\beta = -20.36$, $t = -1.41$	$\beta = -22.96$, $t = -1.00$
Unpred. × NP	$\beta = -0.12$, $z = -0.14$	$\beta = 0.46$, $z = 0.61$	$\beta = -4.17$, $t = -0.25$	$\beta = -7.98$, $t = -0.38$

Notes: Significant effects are in bold. "Unpred." refers to the Unpredictable Context, and the × symbol indicates interactions between factors.

therefore be more marked for adjectives than for nouns, allowing it to serve a focus-marking function. This difference in the distribution of pitch accents for focussed nouns and adjectives may reflect their relative position in the phrase: whereas NP-initial focussed nouns are realised with rising pitch accents, NP-final focussed adjectives are realised with falling pitch accents. Unlike in English, context condition was not a significant predictor of pitch accenting.

Focus condition also affected noun and adjective duration in Guaraní. As shown in the left panel of Figure 6, nouns were significantly shorter in Adjective Focus than Noun Focus, but the duration of nouns in Noun Focus and NP Focus did not differ significantly. These results are consistent with lengthening of focussed expressions in Guaraní (Clopper & Tonhauser, 2013). As shown in the right panel of Figure 6, adjectives were significantly longer in NP Focus than Adjective Focus. Together with

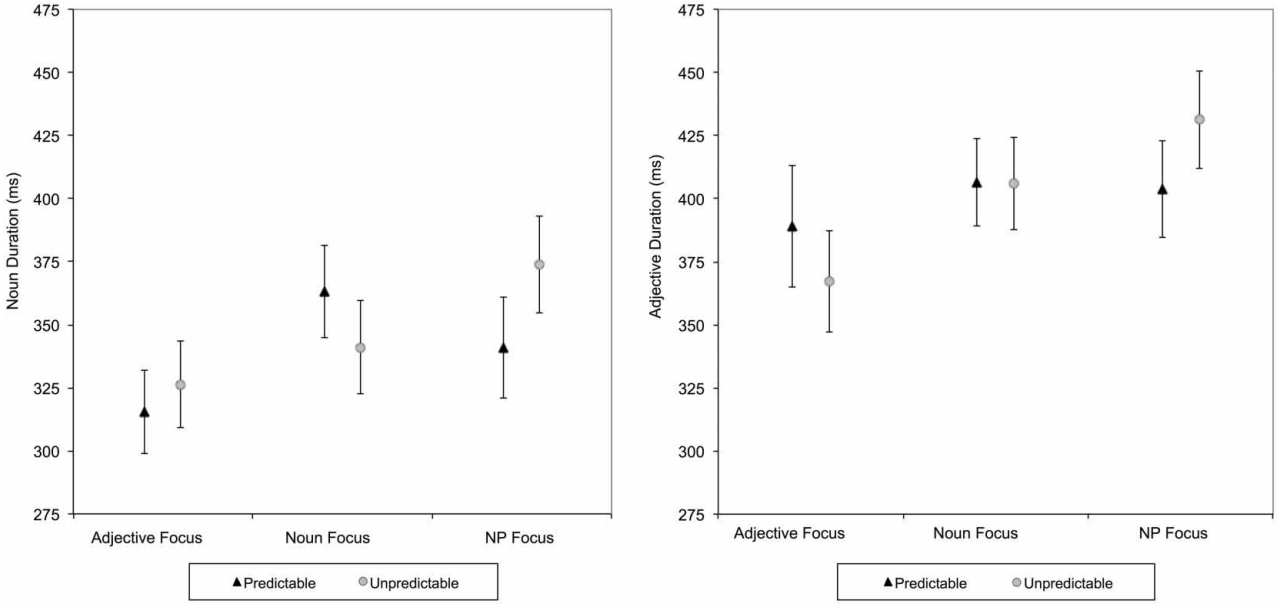

Figure 6. Mean noun (left) and adjective (right) duration in Guaraní for each focus condition in each context condition. Error bars are standard error of Director means.

the noun duration results, this unexpected effect of focus condition on adjective duration may reflect lengthening of both the noun and the adjective in NP Focus. These effects of focus condition on duration were independent of phonological pitch accenting and phrasing. Both nouns and adjectives were longer when they were accented than when they were unaccented (nouns: $\beta = 49.99$, $t = 6.68$; adjectives: $\beta = 25.04$, $t = 2.72$) and when they were followed by a phrase boundary than when they were not (nouns: $\beta = 146.95$, $t = 11.93$; adjectives: $\beta = 72.32$, $t = 13.22$). Unlike in English, context condition was not a significant predictor of word duration.

In contrast with Clopper and Tonhauser's (2013) findings, focus condition had no significant effect on the f0 slope of Guaraní adjectives. However, the absolute value of the f0 slope of the adjectives was affected by context condition, as shown in the right panel of Figure 7. Steeper slopes were observed in the Unpredictable Context than the Predictable Context for accented adjectives, suggesting an overall increase in prosodic prominence, rather than an increase in the prosodic marking of focus, in the Unpredictable Context relative to the Predictable Context. This context effect is independent of effects of pitch accent type and duration: significantly steeper slopes were observed for rising pitch accents than falling pitch accents overall ($\beta = 40.89$, $t = 3.03$) and for longer adjectives than shorter adjectives overall ($\beta = 183.36$, $t = 2.94$). As shown in the left panel of Figure 7, neither focus nor context condition significantly affected the f0 slope of accented nouns.

Taken together, these results demonstrate significant effects of focus and contextual predictability on pitch accenting and duration in Guaraní. Focussed nouns were more likely to be realised with rising pitch accents, whereas non-focussed nouns were more likely to be unaccented. Focussed nouns were also longer than non-focussed nouns. Focussed adjectives were more likely to be realised with falling pitch accents than non-focussed adjectives. The adjectives in focussed NPs were longer than the adjectives in other focus conditions. Unlike in English, context condition did not affect the magnitude of these focus effects. However, the f0 slope of accented adjectives was steeper in all focus conditions in the Unpredictable Context than the Predictable Context, suggesting an overall increase in prosodic prominence in the Unpredictable Context relative to the Predictable Context. Thus, whereas in English effects of focus and context condition were observed across the phonological and phonetic variables we examined, in Guaraní, the effect of focus emerged for pitch accenting and duration and the effect of context condition emerged for f0 slope.

The results from both languages exhibited some asymmetries across nouns and adjectives in different conditions. In English, the duration-lengthening effects of focus and context condition only affected nouns; in Guaraní, f0 slope effects were only observed for adjectives, and deaccenting patterns were only observed for nouns. Differences in NP word order may explain some of these asymmetries, and may prove a fruitful area for future study. As noted above, in Guaraní the adjective follows the noun, whereas in English it precedes the noun. This difference allows for a comparison of effects of position in the phrase (final vs. non-final) and syntactic headedness (the noun being the head of the NP). In English, headedness

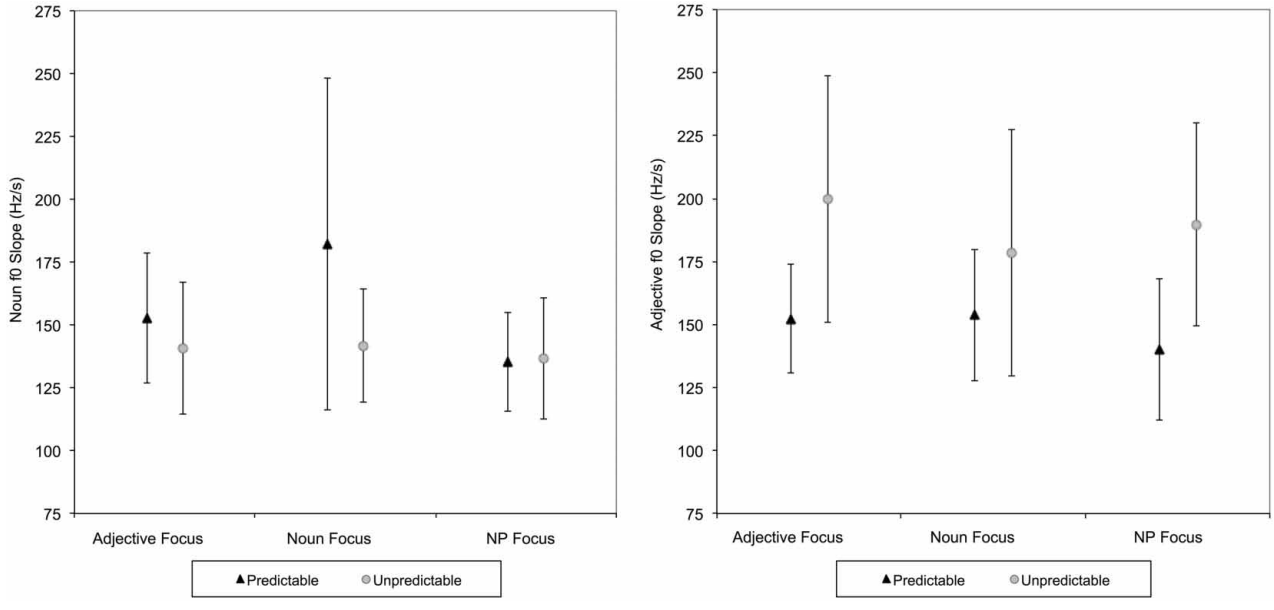

Figure 7. Mean f0 slope for accented nouns (left) and adjectives (right) in Guaraní for each focus condition in each context condition. Error bars are standard error of Director means.

is confounded with finality, but in Guaraní the head noun is non-final. Thus, the f0 slope effects in Guaraní may be due to finality, but the deaccenting patterns to headedness. Our results further demonstrate that Guaraní differs from Romance languages with noun–adjective word order which do not mark focus prosodically within the NP, such as Italian (Ladd, 2008; Swerts et al., 2002), and suggest that word order does not constrain the prosodic realisation of focus within the NP.

General discussion

The same interactive task was used to elicit English and Guaraní utterances with different foci in two conditions that differed in contextual predictability. Our analysis confirmed the prosodic differences between the two languages and further revealed that the prosodic realisation of focus is affected by contextual predictability in both languages. The precise details of these effects, however, appear to be language-specific. In English, the phonetic and phonological cues used to mark focus were enhanced in the Unpredictable Context relative to the Predictable Context, leading to larger prosodic differences between focus conditions in the Unpredictable Context than the Predictable Context. In Guaraní, on the other hand, accented adjectives had steeper f0 slopes in the Unpredictable Context than the Predictable Context regardless of focus condition, suggesting greater levels of overall prosodic prominence in the absence of contextual predictability. Thus, the results of this study revealed cross-linguistic variation in the nature and extent of the effect of contextual predictability of focus on the realisation of prosodic prominence.

The results from both languages also exhibited some asymmetry between the effects observed for nouns and adjectives. As discussed above, we speculate that these differences may be due in part to language-specific word order and syntactic headedness effects. We also obtained one result which deviated from the predicted pattern: in English, a greater proportion of rising pitch accents was observed on the noun in the Predictable Context than the Unpredictable Context. To the extent that rising pitch accents are more prominent than high pitch accents, this result reflects a pattern of increased prominence in the contextually predictable condition relative to the unpredictable condition, which is opposite to the predicted result, and opposite to the pattern of results observed for English adjectives. We do not have an explanation for this deviation from the predicted pattern. Despite these unexpected findings, however, our results provide substantial evidence that the prosodic realisation of focus in both American English and Paraguayan Guaraní is affected by contextual predictability.

Our results thus further support the hypothesis that contextual predictability affects linguistic form at all levels of linguistic structure, including prosody, and extend the empirical support for this hypothesis to a new linguistic domain, the prosodic realisation of focus, and to an understudied language, Guaraní. As described in the introduction, this hypothesis is consistent with information-based theories of language production, such as Aylett and Turk's (2004) smooth signal redundancy hypothesis, Levy and Jaeger's (2007) and Jaeger's (2010) uniform information density hypothesis, and Pluymaekers et al.'s (2005) informational redundancy hypothesis. According to these theories, unpredictable expressions are produced in a phonetically prominent manner to maximise the likelihood of the listener correctly comprehending the utterance. In contrast, predictable expressions are free to be phonetically reduced, thus saving effort for the talker, because their predictable nature means that the listener is unlikely to misunderstand the linguistic content. The English results obtained in this study are generally consistent with these theories, in that the cues marking the focussed expression were larger in the contextually unpredictable condition than in the contextually predictable condition. That is, the English Directors produced reduced prosodic cues to focus in the Predictable Context relative to the Unpredictable Context because the location of focus in their utterances was predictable from the visual context. Similarly, the Guaraní results show that prosodic prominence was enhanced overall in the contextually unpredictable condition. However, the Guaraní pattern was independent of focus marking, suggesting that, unlike in English, the specific cues to focus were not being reduced in the predictable condition. Rather, the overall prominence of the expression was reduced in the Predictable Context relative to the Unpredictable Context through the manipulation of f0 slope. Thus, our results suggest that different languages employ different means to express the effects of contextual predictability on the prosodic realisation of focus and that models of contextual predictability effects must be able to account for the patterns observed in both English and Guaraní in this study.

Additionally, our data do not rule out the possibility that predictability effects on speech production are due to talker-internal cognitive processes, such as contextual information facilitating lexical retrieval and production, rather than information smoothing (Bard et al., 2000; Bell et al., 2003; Gahl, Yao, & Johnson, 2012; Kahn & Arnold, 2012, 2015), or to multiple competing mechanisms, including both information smoothing and egocentric processes (Bard & Aylett, 2005; Watson, 2010). Regardless, the difference in patterns between our English and Guaraní results suggest that any mechanism accounting for these effects must be able to adequately model language-specific differences and their interactions with cognitive processes.

Taken together, our results suggest that the two languages follow the same general principle that prosodic cues are enhanced in the absence of contextual

predictability. However, the specific variables affected by context, and the interaction between contextual predictability and focus marking, is language-specific. In English, cues to focus marking were enhanced in the Unpredictable Context, relative to the Predictable Context. However, this interaction between contextual predictability and focus was not observed in Guaraní; instead, overall prosodic prominence was increased in the Unpredictable Context relative to the Predictable Context, regardless of focus. Thus, our current analyses support the hypothesis that the prosodic realisation of focus is predictability-dependent in English, but that overall prosodic prominence is predictability-dependent in Guaraní. This result underscores the importance of cross-linguistic work in psycholinguistic and prosodic research, and suggests that different languages may respond to contextual manipulations in different ways. Further cross-linguistic research is necessary to determine the possible relationships among contextual predictability and phonological and phonetic cues to focus.

Acknowledgements

The authors thank Sara Phillips-Bourass and Murat Yasavul for assistance with data collection. We are also grateful to our handling editor, Elisabeth Norcliffe, and the anonymous reviewers for *Language, Cognition and Neuroscience* for their detailed comments on previous versions of this paper.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This project was supported by the Department of Linguistics at Ohio State University, the Targeted Investment in Excellence from the College of Arts and Humanities at Ohio State University, and the National Science Foundation [BCS-0952571].

Notes

1. On the assumption that focussed expressions are unpredictable because they provide new or contrastive information, our manipulation is essentially a manipulation of whether the unpredictability of an expression is itself predictable. This kind of predictability of focus from the visual context can be thought of as a "second-order predictability" – the relative expectedness of the unexpected.
2. The Predictable and Unpredictable Contexts differed in the extent to which target words were repeated across utterances within a trial. Whereas non-focussed expressions in Noun and Adjective Focus were repeated four times in the Predictable Context, they were repeated only twice in the Unpredictable Context. If more repetition leads to greater reduction of non-focussed expressions, larger differences between focussed and non-focussed expressions are expected in the Predictable Context than the Unpredictable Context, contrary to our prediction about contextual predictability. Thus, our design may underestimate the effect of contextual predictability on the prosodic realisation of focus.
3. This design involved two levels of contextual predictability. In the Predictable Context, which expression was focussed was predictable from the visual context. In addition, the words in the NP were partially predictable. For example, in the Adjective Focus trial shown in Figure 1, the word for "pig" was predictable from the visual context. Similarly, in NP Focus trials, the adjective and the noun were predictable from each other, since each tile was unique in both shape and colour. In the Unpredictable Context, which expression was focussed was never predictable from the context and the non-focussed expression in Adjective and Noun Focus targets was also not predictable. However, for NP Focus targets, the adjective and the noun were predictable from each other, as in the Predictable Context, because these tiles were unique in both shape and colour among the available tiles.
4. The accent type and focus condition predictors were not significantly correlated with each other in either of the English f0 peak analyses (all $r < .12$).

References

Aissen, J. (2003). Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory*, *21*, 435–483.

Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech*, *47*, 31–56.

Baese-Berk, M., & Goldrick, M. (2009). Mechanisms of interaction in speech production. *Language and Cognitive Processes*, *24*, 527–554.

Bard, E. G., Anderson, A. H., Sotillo, C., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language*, *42*, 1–22.

Bard, E. G., & Aylett, M. (2005). Referential form, word duration, and modeling the listener in spoken dialogue. In J. C. Trueswell & M. K. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions* (pp. 173–192). Cambridge, MA: MIT Press.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278.

Beckman, M. E., & Ayers Elam, G. (1997). *Guidelines for ToBI labelling (version 3.0)*. Retrieved from http://www.ling.ohio-state.edu/~tobi/ame_tobi/labelling_guide_v3.pdf

Beckman, M. E., Hirschberg, J., & Shattuck-Hufnagel, S. (2005). The original ToBI system and the evolution of the ToBI framework. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 9–54). Oxford: Oxford University Press.

Bell, A., Brenier, J. M., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, *60*, 92–111.

Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., & Gregory, M. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *The Journal of the Acoustical Society of America*, *113*, 1001–1024.

Bolinger, D. (1972). Accent is predictable (if you're a mind-reader). *Language*, *48*, 633–644.

Braun, B. (2006). Phonetics and phonology of thematic contrast in German. *Language and Speech*, *49*, 451–493.

Breen, M., Fedorenko, E., Wagner, M., & Gibson, E. (2010). Acoustic correlates of information structure. *Language and Cognitive Processes*, *25*, 1044–1098.

Burdin, R. S., Phillips-Bourass, S., Turnbull, R., Yasavul, M., Clopper, C. G., & Tonhauser, J. (in press). Variation in the prosody of contrastive focus in head- and edge-marking languages. *Lingua*. doi:10.1016/j.lingua.2014.10.001

Calhoun, S. (2010). The centrality of metrical structure in signaling information structure: A probabilistic perspective. *Language*, *86*, 1–42.

Charles-Luce, J. (1993). The effects of semantic context on voicing neutralization. *Phonetica*, *50*, 28–43.

Clopper, C. G., & Pierrehumbert, J. (2008). Effects of semantic predictability and regional dialect on vowel space reduction. *The Journal of the Acoustical Society of America*, *124*, 1682–1688.

Clopper, C. G., & Tonhauser, J. (2013). The prosody of focus in Paraguayan Guaraní. *International Journal of American Linguistics*, *79*, 219–251.

Eady, S. J., & Cooper, W. E. (1986). Speech intonation and focus location in matched statements and questions. *The Journal of the Acoustical Society of America*, *80*, 402–415.

Féry, C., & Kügler, F. (2008). Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics*, *36*, 680–703.

Fowler, C. A., Levy, E. T., & Brown, J. M. (1997). Reductions of spoken words in certain discourse contexts. *Journal of Memory and Language*, *37*, 24–40.

Gahl, S. (2008). Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language*, *84*, 474–496.

Gahl, S., Yao, Y., & Johnson, K. (2012). Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *Journal of Memory and Language*, *66*, 789–806.

Gregores, E., & Suárez, J. (1967). *A description of colloquial Guaraní*. The Hague: Mouton de Gruyter.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, *69*, 274–307.

Ito, K., & Speer, S. R. (2006). Using interactive tasks to elicit natural dialogue. In S. Sudhoff, D. Lenertová, R. Meyer, S. Pappert, P. Augurzky, I. Mleink, N. Richter, & J. Schließer (Eds.), *Methods in empirical prosody research* (pp. 227–257). Berlin: de Gruyter.

Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology*, *61*, 23–62.

Jun, S.-A. (2014). Prosodic typology: By prominence type, word prosody, and macro-rhythm. In S.-A. Jun (Ed.), *Prosodic typology II: The phonology of phrasing and intonation* (pp. 520–540). Oxford: Oxford University Press.

Kahn, J., & Arnold, J. E. (2012). A processing-centered look at the contribution of givenness to durational reduction. *Journal of Memory and Language*, *67*, 311–325.

Kahn, J., & Arnold, J. E. (2015). Articulatory and lexical repetition effects on durational reduction: Speaker experience vs. common ground. *Language, Cognition and Neuroscience*, *30*, 103–119.

Katz, J., & Selkirk, L. (2011). Contrastive focus vs. discourse-new: Evidence from phonetic prominence in English. *Language*, *87*, 771–816.

Krahmer, E. J., & Swerts, M. (2001). On the alleged existence of contrastive accents. *Speech Communication*, *34*, 391–405.

Ladd, D. R. (1980). *The structure of intonational meaning: Evidence from English*. Bloomington, IN: Indiana University Press.

Ladd, D. R. (2008). *Intonational phonology*. Cambridge: Cambridge University Press.

Leonetti, M. (2004). Specificity and differential object marking in Spanish. *Catalan Journal of Linguistics*, *3*, 75–114.

Levy, R., & Jaeger, T. F. (2007, December). *Speakers optimize information density through syntactic reduction*. Proceedings of the 20th annual conference on neural information processing systems, Vancouver, BC.

Myers, J., & Li, Y. (2009). Lexical frequency effects in Taiwan Southern Min syllable contraction. *Journal of Phonetics*, *37*, 212–230.

Pierrehumbert, J., & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. Cohen, J. Morgan, & M. Pollack (Eds.), *Intentions in communication* (pp. 271–311). Cambridge, MA: MIT Press.

Pitrelli, J. F., Beckman, M. E., & Hirschberg, J. (1994, September). *Evaluation of prosodic transcription labeling reliability in the ToBI framework*. Proceedings of the international conference on spoken language processing, Yokohama, Japan, pp. 123–126.

Pluymaekers, M., Ernestus, M., & Baayen, R. H. (2005). Articulatory planning is continuous and sensitive to informational redundancy. *Phonetica*, *62*, 146–159.

Raymond, W. D., Dautricourt, R., & Hume, E. (2006). Word-internal /t, d/ deletion in spontaneous speech: Modeling the effects of extra-linguistic, lexical, and phonological factors. *Language Variation and Change*, *18*, 55–97.

Riester, A., & Piontek, J. (in press). Anarchy in the NP. When new nouns get deaccented and given nouns don't. *Lingua*. doi:10.1016/j.lingua.2015.03.006

Roberts, C. (2012). Information structure in discourse. Towards an integrated formal theory of pragmatics. *Semantics & Pragmatics*, *5*(6), 1–69.

Selkirk, L. (1995). Sentence prosody: Intonation, stress, and phrasing. In J. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 550–569). 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*, 103–130.

Speer, S. R., Warren, P., & Schafer, A. J. (2011). Situationally independent prosodic phrasing. *Laboratory Phonology*, *2*, 35–98.

Swerts, M., Krahmer, E., & Avesani, C. (2002). Prosodic marking of information status in Dutch and Italian: A comparative analysis. *Journal of Phonetics*, *30*, 629–654.

Syrdal, A. K., & McGory, J. T. (2000, October). *Inter-transcriber reliability of ToBI prosodic labeling*. Proceedings of Interspeech 2000, Beijing, China.

Tily, H., & Kuperman, V. (2012). Rational phonological lengthening in spoken Dutch. *The Journal of the Acoustical Society of America*, *132*, 3935–3940.

Vajrabhaya, P., & Kapatsinski, V. (2011, August). *There is more to the story: First-mention lengthening in Thai interactive discourse*. Proceedings of the international congress of phonetic sciences XVII, Hong Kong.

Van Son, R., Bolotova, O., Lennes, M., & Pols, L. C. W. (2004, October). *Frequency effects on vowel reduction in three typologically different languages (Dutch, Finnish, Russian)*. Proceedings of Interspeech 2004, Jeju Island, Korea.

Van Son, R., & Pols, L. C. W. (2003, September). *Information structure and efficiency in speech production*. Proceedings of Interspeech 2003, Geneva, Switzerland.

Velázquez-Castillo, M. (2004). Guaraní (Tupí-Guaraní). In G. Booij & C. Lehmann (Eds.), *Morphology: An international handbook on inflection and word formation* (pp. 1421–1432). Berlin: Walter de Gruyter.

Vigário, M. (2003). *The prosodic word in European Portuguese*. Berlin: Mouton de Gruyter.

Watson, D. G. (2010). The many roads to prominence: Understanding emphasis in conversation. *Psychology of Learning and Motivation*, *52*, 163–183.

Watson, D. G., Arnold, J. E., & Tanenhaus, M. K. (2008). Tic tac TOE: Effects of predictability and importance on acoustic prominence in language production. *Cognition*, *106*, 1548–1557.