Thorson, J. (2007). The scaling of utterance-initial pitch peaks in Puerto Rican Spanish: Evidence for tonal preplanning. In L. Wolter and J. Thorson (Eds.), *University of Rochester Working Papers in the Language Sciences, 3(1),* 91-97.

THE SCALING OF UTTERANCE-INITIAL PITCH PEAKS IN PUERTO RICAN SPANISH: EVIDENCE FOR TONAL PREPLANNING

Jill Thorson
University of Rochester

For this study, an experiment was created to test for the existence of tonal preplanning in Puerto Rican Spanish. The scaling of utterance-initial F0 high peaks was examined as a function of utterance length, which was defined as total utterance duration (in seconds). Eight speakers from Caguas, Puerto Rico, were recorded reading 24 distinct utterances of varying length for analysis. The data revealed that all eight of the subjects began higher in their pitch range as utterance length increased. Results show that tonal preplanning is present and that the *lookahead* mechanism is being employed. Due to variation within and between speakers and the presence of other tonal contour events such as downstep and upstep, the conclusions accord with recent studies in Romance tonal preplanning that label this phenomenon as *soft* preplanning (Liberman & Pierrehumbert 1984, Prieto et. al. 2006).

1. Introduction

The goal of this paper is to examine the phenomena of preplanning in the tonal production of Puerto Rican Spanish through the investigation of the scaling of utterance-initial F0 peaks.

1.1 Background and Current Work

Tonal preplanning refers to a *lookahead* strategy by which a speaker will vary the initial F0 values of an utterance depending on how long the utterance is, by setting the initial F0 value higher for longer utterances. The degree to which speakers utilize such a *lookahead* mechanism has been heavily debated (Liberman & Pierrehumbert 1984, Prieto & Shih 1996, Prieto et. al. 2006, Thorsen 1980, among others). Primarily, debate has centered around whether pitch is employed by speakers on a global or local level. The *global hypothesis* is the belief that utterance length alone determines the initial F0 values of a speaker's pitch contour; a speaker will anticipate the length of an utterance before beginning production. In the *local hypothesis*, tonal decisions are made locally during the production of the utterance (Prieto et. al. 2006). The local hypothesis predicts that utterance length does not set utterance-initial pitch values, but rather speakers will make pitch decisions on an accent-by-accent basis.

The degree to which preplanning is implemented, and the *lookahead* mechanism used by speakers, can be thought to exist along a continuum. The edges of this continuum were named *hard* and *soft* preplanning by Liberman and Pierrehumbert (Liberman & Pierrehumbert 1984). In their view, hard preplanning has two primary requirements. First, it is a global phenomenon where utterance-initial F0 values are determined exclusively by the length of the utterance. Second, it is not an optional mechanism; its tonal realization is required by *all* speakers of a language in all circumstances. *Hard* preplanning, in this view, is an "essential part of intending to say something" and, in order to be fluent, is accomplished before the execution of this intention (Liberman & Pierrehumbert 1984). The other end of the spectrum is *soft* preplanning, which refers to a local *lookahead* strategy. It is not a mandatory feature of tonal production and its optional implementation may vary by speaker and/or by language. One prediction of this approach is that speakers may run out of range or hit asymptote before the end of the utterance.

The *lookahead* problem of advanced planning is briefly mentioned in *Intonational Phonology* by Ladd, "part of the reason it is difficult to provide quantitative descriptions of F0 trends is that in general such trends are quite plainly dependent on the length of the domain to which they apply" (Ladd 1996). Ladd goes on to state that a lookahead mechanism "may be psycholinguistically implausible," but he does not make any reference to utterance-initial pitch, only to declination and the outcome of the slope over time (Ladd 1996).

In attempting to pull apart these issues, current studies have addressed preplanning more directly. Hansson discusses that "hard" preplanning refers to articulatory processes that *need* to be mapped out before a phrase begins, while 'soft' preplanning refers to *optional* production strategies that the speaker may make before beginning" (Hansson 2003). Hansson states that "an obligatory adjustment of the F0 starting point to phrase length, as suggested in the early work on downstep by Bruce (1982a), would also require some lookahead and hard preplanning" (Hansson 2003). However, in her work on spontaneous Swedish speech, she found that speakers are not using *hard* tonal preplanning in spontaneous speech and instead presents data that they are making decisions on a phrase-by-phrase basis, (i.e. *soft* preplanning).

A recent study by Prieto et. al. has also shown evidence in support of *soft* preplanning in several of the Romance languages: Central Catalan, Central Spanish, Neapolitan Italian, Northern European Portuguese and Standard European Portuguese (Prieto et. al. 2006). This study examined the initial pitch of an utterance and the pitch of the first peak of an utterance in order to determine if the initial pitch became higher with increasing utterance length. One important finding was that pitch of the first F0 peak more clearly reflected the presence of tonal preplanning than the initial pitch of the utterance when correlated with utterance length. They concluded that it was *soft* preplanning that existed and not *hard* preplanning due to the fact that, although it was implemented by the majority of speakers, not *all* of their speakers exhibited a correlation between initial F0 values and utterance length.

This paper will examine the scaling of utterance-initial high pitch peaks, as opposed to initial F0 values, as a function of utterance length in order to determine if *soft* or *hard* preplanning is being employed by the speakers.

2. Method

A set of sentences was read aloud by the experiment participants at a normal speaking rate. The pitch of the initial F0 peaks was then measured and plotted against the total duration of the utterance (in seconds).

2.1 Test Utterances

Two sets of test utterances were designed as part of a larger study on the intonation of Puerto Rican Spanish. Each set consisted of twelve declarative utterances of varying length, giving a total of 24 distinct utterances.

Two factors were manipulated in each utterance to vary length: the number of pitch accents (from 1 to 5) and the number of intervening unstressed syllables between pitch accents (from 0 to 6). The utterances as a whole varied in total number of syllables from three to fourteen. Examples of utterances of varying number of pitch accents are listed below in Table 1.

Table 1: Example utterances by number of pitch accents. The number of intervening unstressed syllables varies between pitch accents.

# of Accents	Example/Translation
1	*Le amaron.* They loved him/her.
2	*Amaba la mina.* He/She loved the mine.
3	*María mira al niño.* Maria looked at the child.
4	*Un año sin lo del amor ha pasado.* A year without love has passed.
5	*La mala niña de la fiesta está aquî.* The naughty girl from the party is here.

2.2 Participants and Procedures

Four female and four male native speakers of Puerto Rican Spanish participated in the experiment. The speakers ranged in age from 13 to 60, with an average age of 22. All eight of the speakers were from and were recorded in Caguas, Puerto Rico.

The participants read aloud two sets of utterances three times, after having the opportunity to practice using a pre-test set of utterances. Half of the participants read the sets in reverse order. The second repetition of each utterance was used for analysis (8 speakers x 24 utterances = 192 utterances).

2.3 Recordings and Data Analysis

Speech recordings were made on a digital Marantz recorder using a Shure head-mounted microphone and sampled at 44.1 KHz. The recordings were then transferred to computer by using a digital-to-digital USB cable and each utterance was saved in individual .wav files. Speech and pitch analysis was conducted using the Praat software for Macintosh using standard settings (Boersma & Weenink).

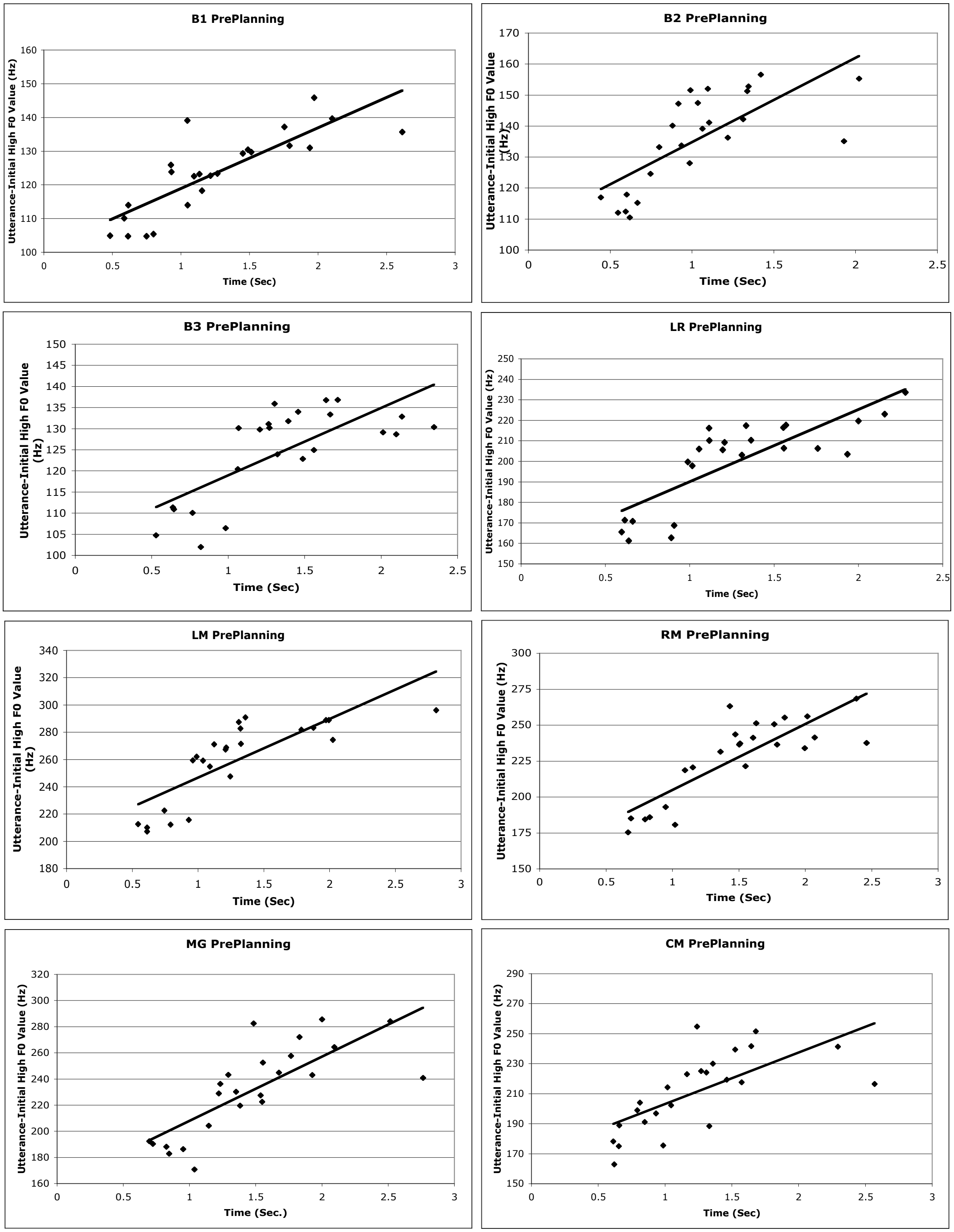
A Praat script was written to extract the initial F0 peak values then place them into an Excel file for analysis along with the utterance text and the utterance duration (in seconds). Values for F0 were double-checked manually for accuracy to adjust for tracking errors.
Utterance length was originally analyzed in three ways for this study: by number of syllables, by number of pitch accents and by total utterance duration.

3. Results

3.1 Scaling of Utterance-Initial F0 Peaks

For this paper, utterance length was analyzed as the total duration of the utterance (in seconds). The F0 values of the initial peaks were measured to examine whether or not they fluctuated significantly as a function of utterance length. Figure 1 on the following page shows graphs that plot utterance-initial pitch peak height (in Hz) against utterance duration (in sec.).

Figure 1: Utterance-initial high peak values (in Hz) on y-axis; total utterance duration in seconds on x-axis. The solid black line is the linear regression for each plot. Each graph is a different speaker, the top four are male and the bottom four are female.



For every speaker, the graphs in Fig. 1 clearly show a positive trend line (the linear regression), which displays a positive correlation between utterance-initial F0 peak height (in Hz) and utterance length (in sec.). The slopes of the regression lines vary by graph, with some more steep than others.

Also, in order to test for variability within the data and total scatter of the plotted points, the R squared values were found for each graph. Table 2, below, lists the R squared values along with the slopes of the regression lines for each subject. All the slopes are positive.

Table 2: This table gives the slope and the R squared values for the linear regression lines of each subject.

Subject	Slope	R squared
Boy1	18.01	0.67
Boy2	27.18	0.52
Boy3	15.97	0.53
LR	35.21	0.65
MG	49.11	0.59
RM	45.85	0.69
LM	43.07	0.64
CM	34.26	0.45

4. Discussion

4.1 Utterance Length

For this paper, only the data for utterance length as duration in seconds is presented. Number of pitch accents was omitted due to the fact that the number of intervening unstressed syllables did not remain constant across the utterances. This variation created utterances of different lengths even when they contained the same number of pitch accents.

Utterance length as number of syllables was omitted in favor of total duration (in seconds). Even though Spanish is a syllable-timed language and has a relatively low PVI (Pairwise Variability Index) for syllable length, the varying onset and coda lengths and stressed versus unstressed vowel durations result in uneven syllable durations among the tokens (Grabe & Low 2002, Ladefoged 2003). Thus, utterance length as total temporal duration was chosen to reflect the most accurate account of tonal preplanning.

4.2 Soft Preplanning

For the results reported in this paper, all of the eight speakers recorded show positive regression lines. The upward trend indicates that as an utterance increases in duration so does its initial pitch peak (in Hz), implying the presence of a *lookahead* mechanism being employed by all eight subjects and suggesting that we are seeing a global phenomenon. However, an alternate analysis may be provided.

At first glance, these data look to be supporting *hard* preplanning because all speakers exhibit a positive correlation with utterance length. But, as Liberman and Pierrehumbert note, a global phenomenon may be described by a series of local ones. For example, they describe overall utterance declination in pitch as a series of downstepped contours, with final lowering at the end of the utterance, all local events (Liberman & Pierrehumbert 1984). Similarly, although it may appear that the data from Puerto Rican Spanish resembles a global output, the pieces at work may also be best understood as operating locally.

Evidence that a local level over a global phenomenon is at work lies in the variation in the data. The overall positive regression lines do not take into account the variation within and between subjects. Table 2 in the results section reports the varying steepness of the slopes and the R squared values that represent the overall variation in the points. R squared values range from .45 to .69, implying variation among the subjects. In fact, as is visible from the graphs in Figure 1, individual points do vary, sometimes extensively. It is difficult to say which preplanning hypothesis this kind of variability supports. However, *soft* preplanning does not preclude positive regression and it predicts looser alignment F0 to the more local events, such as utterance initial peaks.

In addition to the variation of individual data points, there are other events that occur in the utterances that are not well represented by utterance-initial pitch presetting. Though not reported on in this study, upstep and downstep were also present in the utterances analyzed; these behaviors lend themselves to a local interpretation as speakers reset their pitches against local pitch events.

Finally, the data from this study accords with the Central Spanish speaking portion of the Prieto et. al. study where, interestingly, both speakers implement tonal preplanning. They state that their data implies a global preplanning phenomenon but end up concluding with a *soft* preplanning analysis since not every speaker uses it across all of the languages analyzed (i.e. Standard European Portuguese) (Prieto et. al. 2006). I suggest that even within a population with a significant correlation between utterance length and initial pitch, *soft* preplanning is present.

5. Conclusions

The findings of this study show a strong trend for speakers to begin with higher utterance-initial F0 peaks in longer utterances. This is representative of a global preplanning strategy. A *lookahead* mechanism is being employed by the speakers in this experiment. Although it can be argued that this is evidence of *hard* preplanning, several issues arise that contest this. Variability in implementation, local tonal events within the pitch trace (i.e. upstep) and the question of the definition of locality, along with previous studies, support a *soft* preplanning approach over *hard* preplanning. This study further suggests that a distinction between *hard* and *soft* preplanning is likely to be gradient. Neither extreme of the preplanning continuum is likely to exist due to the variability of human speech, as well as factors external to a physical speech event, such as speaker's intent. Instead, investigation of the role of variability in preplanning and an articulated definition of locality is warranted.

Future work might analyze utterances with and without upstep and downstep in order to examine how these affect tonal preplanning.

References

Boersma, P., Weenink, D. Praat: doing phonetics by computer (Version 4.4.20) [Computer program]. http://www.praat.org/ Retrieved 1-Aug-06

Grabe, E., Low, EL. (2002). Durational variability in speech and the rhythm class hypothesis. In: Gussenhoven, C., Warner, N. (eds), *Papers in Laboratory Phonology 7*, Mouton, 515-546.

Hansson, P. (2003). Chapter 5: Tonal coherence among prosodic phrases. *Prosodic Phrasing in Spontaneous Swedish*. Lunds Universitet, Sweden.

Ladd, R. 1996. *Intonational Phonology*. Cambridge University Press.

Ladefoged, P. (2003). *A Course in Phonetics*. 5th ed. Boston: Thomson Wadsworth.

Liberman, M., Pierrehumbert, J. (1984). Intonational invariance under changes in pitch range and length. *Language and Sound Structure*. Cambridge, MA: MIT Press, 157-233.

Prieto, P., Shih, C. (1996). Pitch downtrend in Spanish. J*ournal of Phonetics* 24, 445-473.

Prieto, P., D'Imperio, M., Elordieta, G., Frota, S., Vigáro, M. (2006). Evidence for *soft* preplanning in tonal production: Initial scaling in Romance. *Proc. Speech Prosody Conference* Dresden.

Thorsen, N. (1980). Intonation contours and stress group patterns in declarative sentences of varying length in ASC Danish. *Annual report of the Institute of Phonetics*, University of Copenhagen. 14, 1-29.

Acknowledgements

I would like to acknowledge and thank my adviser, Prof. Joyce McDonough, for all of her help, knowledge, and understanding.

Linguistics Department
University of Rochester
Rochester, NY 14627

thorson@ling.rochester.edu

UNIVERSITY OF ROCHESTER WORKING PAPERS IN THE LANGUAGE SCIENCES — VOL. 3, NO. 1

Lynsey Wolter and Jill Thorson, Editors
Joyce McDonough, Series Editor