Speech Prosody 2016
31 May - 3 Jun 2106, Boston, USA

Is the input for prosodic bootstrapping of word order reliable? The case of phrasal prominence in Turkish and French

Angeliki Athanasopoulou[1]*, Irene Vogel*[1]

[1] University of Delaware
angeliki@udel.edu, ivogel@udel.edu

Abstract

Prosody is often attributed a fundamental role in the process of language acquisition, allowing infants to use prosodic cues to begin to acquire the syntactic structures of their language. The Prosodic Bootstrapping process may combine a number of phonological phenomena (e.g., stress, rhythmic units, intonation), and recently, it has been proposed that a Rhythmic Activation Principle integrating the Iambic-Trochaic Law also contributes to bootstrapping, with the acoustic characteristics of phrasal prominence cuing a language's basic word order [1, 2]. That is, association of the different properties of iambic and trochaic stress with phrasal prominence patterns would allow a child to identify whether a language is syntactically VO or OV (see [1, 2] for French and Turkish). We further test this hypothesis by investigating the acoustic manifestations of stress in Turkish and French, specifically by comparing them in non-focus and focus conditions.

Index Terms: Prosodic Bootstrapping, Rhythmic Activation Principle, Complement Law, Iambic/Trochaic Law, Focus.

1. Introduction

The Prosodic Bootstrapping Hypothesis (PBH) proposes that pre-linguistic infants use prosodic cues in speech to learn syntactic structures via perceptual filters and attentional biases [3]. A more targeted aspect of PBH, referred to here as the Phrasal Prominence Hypothesis (PPH), additionally draws on a combination of the Iambic/Trochaic Law, a general auditory bias, and the Complement Law as used in the Rhythmic Activation Principle [1, 2]. We test the feasibility of the Phrasal Prominence Hypothesis, specifically the claim that infants use the prominence patterns of (Phonological) Phrases to learn the basic word order of their language, by comparing the acoustic properties of words in non-focus vs. focus contexts in Turkish (OV) and French (VO).

2. Prosodic Bootstrapping

2.1. The Prosodic Bootstrapping Hypothesis (PBH)

The PBH attempts to answer one of the main questions in language acquisition research: how preverbal infants develop a grammar from the input speech. The proposal is that various acoustic properties present in the input (e.g., stress, rhythm, intonation, phonotactics) provide infants with crucial cues to determine basic aspects of the grammatical organization of their language [4, 5, 6]. Two core assumptions [3, 5, 6] underlying the PBH are thus: a) syntactic and prosodic units are reliably correlated with each other, and b) preverbal infants are sensitive to and able to use the acoustic characteristics of the prosodic units. In the present paper, we share assumption b and investigate a – essentially the reliability of acoustic cues to signal the information necessary for bootstrapping.

2.2. PBH and Rhythmic Activation Principle (RAP)

While boundary phenomena (e.g., final lengthening, pitch movements, pausing) are proposed as cues to large syntactic units (i.e., phrases, clauses), the RAP [7, 8] further proposes that the Strong (S) / Weak (W) rhythmic properties of a language serve as cues to phrase and sentence internal structure, specifically basic word order. That is, according to the Complement Law (CL), (syntactic) heads of phrases are weak with respect to their (syntactic) complements, so when a baby hears a SW or WS rhythmic pattern, this informs him/her of the corresponding word order.

Crucially, a WS rhythm will signal that the head of a phrase precedes its complement, as in French, where the order in a (verb) phrase is verb-object (VO). A SW rhythm will signal instead that the head of a phrase follows its complement, as in Turkish, where the order in a (verb) phrase is object-verb (OV). At the next level, the Intonational Phrase (IP), a PP with narrow focus is Strong; otherwise the default Strong is on the final PP, regardless of basic word order.

2.3. PBH and the Iambic-Trochaic Law (ITL)

The ITL was initially proposed to explain perceptual patterns in non-linguistic rhythmic groupings [9, 10, 11, 12]. Sequences of sounds alternating in intensity are perceived with the more prominent (i.e., louder) sound at the beginning of each group, while sequences alternating in duration are perceived with the more prominent (i.e., longer) sound at the end of each group. This extralinguistic principle was then introduced into linguistic theory to account for word-level prominence patterns in terms of foot structure [13, 14, 15]. Sequences of syllables differing in intensity form trochaic feet, while those differing in duration form iambic feet, the latter also commonly considered quantity-sensitive.

In the extension of the ITL to the phrasal level, pitch is incorporated along with intensity as a cue to a SW (trochaic) pattern between words (though there may also be some duration differences). If, instead, words differ in duration (and optionally in intensity), they form iambic structures [1, 2]. The correspondence between word and phrase level in iambic and trochaic structures can be seen in the grid representations in (1), with examples from [1, 2]. At the phrase level, the "o" and "X" indicate the weak and strong positions, respectively. As can be seen, French and Turkish are identical up to the word level (both having final stress), but they crucially differ at the phrase level, having WS and SW patterns, respectively. We

 doi: 10.21437/SpeechProsody.2016-141

refer to the combination of the ITL with the CL as the Phrasal Prominence Hypothesis (PPH).

(1) *Iambic vs. Trochaic Phrasal Structure*

	French (Iambic)	**Turkish (Trochaic)**
Phrase	(o X)	(X o)
Word	(x) (x)	(x) (x)
Foot	(x) (x)	(x) (x)
Syllable	x x x x	x x x x
	antique kilim	kilim için
	'ancient kilim'	'kilim for'

2.4. PBH and Phrasal Prominence Hypothesis (PPH)

The evidence for the PPH comes primarily from French (VO) and Turkish (OV), both of which have fixed word-stress on the final syllable; the data were also compared to VO and OV structures in German [1, 2]. Given the different use of prominence properties in WS (iambic) and SW (trochaic) structures, it was predicted that French would manifest "S" with greater vowel duration while Turkish would use pitch and intensity. In fact, it was found that in addition to these properties, French also made some use of intensity, and Turkish also made some use of duration.

Across languages, it was further expected that the (strong) stressed vowel would have higher pitch and intensity but shorter duration in Turkish than in French. Instead, Turkish exhibited greater values for all measures [1, 2]. Given that the S vowels were phrase final in French, it is possible that their F0 and intensity also incorporated pre-boundary lowering. No additional analysis was provided, however, to determine the relative roles of the significant cues or of the boundaries. Moreover, although the PPH is proposed to apply to the broad focus context, the carriers placed the targets in narrow focus. That is, the right edge of the target phrases aligned with a strong prosodic boundary (henceforth, Intonational Phrase (IP) boundary), not just a PP boundary, resulting in a confound: the Strong word in French has the additional IP boundary properties, while it is the Weak word in Turkish that has the additional properties.

Finally, a language-specific "accent-initial" pattern was observed in French, raising pitch on the Weak word. Since this is not combined with increased intensity as in Turkish, it is not necessarily problematic for the PPH, but it does raise a question for bootstrapping since it potentially introduces a contradiction for the pre-linguistic infant hearing a phrase with pitch properties that indicate initial prominence but duration properties that indicate final prominence.

3. Present Study - Design

3.1. Hypothesis: Word Order and Acoustic Cues

Given the PBH's crucial assumption of the reliability of acoustic cues to signal syntactic structures, conflicting indicators, as described, raise concerns about their feasibility in the process of bootstrapping, in this case, the determination of basic word order. Since there was a potential confound due to the presence of an IP boundary at the right edge of the stimuli, we examine the manifestation of prominence in the same languages in a way that avoids this confound. Thus, we test the same basic hypothesis that underlies the PPH regarding prominence cues in VO and OV languages. The hypothesis is specifically formulated as in (2).

(2) *Word Order – Acoustic Cues Hypothesis*

In French, Strong vs. Weak words in a phrase are distinguished by duration; in Turkish, they are distinguished by pitch and intensity.

3.2. Participants and Procedure

The participants were 15 university educated native speakers of Turkish and French (age: 20-25), recorded in Istanbul and Paris, respectively. They were trained and practiced first, and then read a series of two-part dialogues in a PowerPoint presentation on a computer screen. The dialogues alternated with filler slides showing familiar objects that the participants had to name in order to reduce repetitive prosodic patterns. The productions were recorded in a quiet room onto a computer using an external head-mounted microphone.

3.3. Stimuli

While our stimuli were originally designed for another study, they also provide the opportunity to test the PPH. For both languages, we used real 3-syllable words containing 10 each of /i, o, a/ in the (final) stressed syllable in French, and /i u, a/ in Turkish (e.g., French: *syndicat* [sɛ̃dika] 'union'; Turkish: *karada* [karada] 'on the land').

In order to compare the manifestation of prominence in prosodically Strong vs. Weak positions (comparable to the complement and head positions in phrases), we examine the properties of the (stressed) target vowels in two contexts. However, instead of comparing them in the same (focused) IP constituent, we systematically vary the S / W status of the targets by eliciting them in two prosodic positions: at a (strong) IP boundary induced by focus, and internal to an IP with focus on a subsequent word. This allows us to use the same prosodic structures for French and Turkish, and thus avoids the confound in which the IP boundary at the end of the target phrase coincided with the (strong) complement position in French, but with the (weak) head in Turkish. The two contexts we compared were elicited by dialogues, illustrated for French and Turkish in (3) and (4). The target is indicated with quotation marks, and the vowel we measured is the one in the response portion; the focused words are in bold here.

(3) *French Target Structures.*

<u>a. Strong Condition (target word at IP boundary):</u>

*Marie a dit «**syndicat**»]$_{STRONG}$ cet après-midi.*
Marie said «**union**» this afternoon
Marie said "**union**"]$_{STRONG}$ this afternoon.

<u>b. Weak Condition (target word IP internal)</u>

*Non. Marie a dit «syndicat»]$_{WEAK}$ cet **après-midi**, pas ce matin.*
No. Marie said «union» this **afternoon**, not this morning
'No. Marie said "union"]$_{WEAK}$ this **afternoon,** not this morning.'

(4) *Turkish Target Structures.*

<u>a. Strong Condition (target word at IP boundary):</u>

*Ayşe öğleden sonra "**karada**"]$_{STRONG}$ dedi.*
Ayse afternoon "**on the land**" said
'Ayse said "**on the land**"]$_{STRONG}$ in the afternoon.'

<u>b. Weak Condition (target word IP internal):</u>

*Hayır. Ayşe öğleden sonra "karada"]$_{WEAK}$ **dedi**, yazmadi.*

No. Ayse afternoon “on the land” **said**, not wrote
‘No. Ayse **said** “on the land”]WEAK in the afternoon, she didn’t write it.’

The grid representations with the targets in the Strong and Weak contexts are shown in (5). In the Strong context, the target aligns with the right edge of the focused constituent, and thus a strong boundary, comparable to the one used in the previous studies. In the Weak context, the focus falls on a word to the right of the target, so the target itself is Weak.

(5) *Strong and Weak Target Positions.*

	Strong Target	**Weak Target**	
Phrase	(X)	(o X)	
Word	(x)	(x)	…
Foot	(x)	(x)	…
Syllable	x x x	x x x	…
<u>French</u>	syn di cat	syn di cat	‘union’
<u>Turkish</u>	ka ra da	ka ra da	‘on the land’

3.4. Analysis

3.4.1. Acoustic Analysis

The data were segmented and analyzed in Praat. The measures analyzed are: duration, mean F0 (excluding first 10ms of the vowel), ΔF0 (change from beginning to end of vowel), mean intensity and vowel centralization. The data were normalized with Z-scores to abstract away from individual speaker and vowel differences and thus permit pooling of data.

3.4.2. Statistical Analysis

In order to investigate the differences between the properties in the Strong and Weak prosodic conditions, Binary Logistic Regression Analyses (BLRAs) were run for each language. These allow us to determine not only which properties are significantly different in the two conditions, but also to what extent each variable predicts the distinction (classification) of the target vowels in these conditions.

4. Results

While it can be expected that for each language the Strong and Weak targets will be distinguishable, the crucial information here is whether the main properties in the classification are Duration for French and Pitch and Intensity for Turkish.

4.1. French

In French, overall, the Strong and Weak words were correctly classified 89% of the time by the BLRA, using all of the acoustic measures. As Table 2 shows, while Duration was a significant predictor, it only distinguished 63% of the targets. By contrast, the Pitch properties (ΔF0 and mean F0) were much more successful classifiers (83% - 88%).

Table 1. *BLRA results for French.*

Overall Classification	**Individual Predictors**	**Classification for Individual Predictors**
89%	▪ ΔF0 ▪ Mean F0 ▪ Duration	88% 83% 63%

To understand the acoustic patterns, we also examined the pooled z-scores for the different properties, shown in Figure 1.

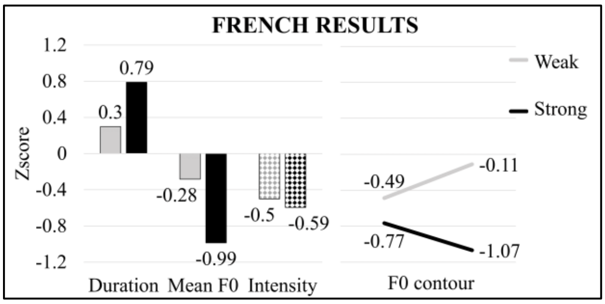

Figure 1: *French - Duration, F0, Intensity and Contour Results. Textured pattern indicates non-significant property.*

As can be seen, Duration, which according to the PPH should be the main cue for prominence in French, does show an increase in the Strong items, but as mentioned, the rate of classification is not very high. The strongest classifier, ΔF0, reveals a clear falling contour in the S items in contrast with a rising contour in W ones. Mean F0 is also a strong predictor, though it is lower in the Strong items than the Weak items, whereas it would be expected that the Strong items would exhibit a higher F0. As expected, the values for Intensity are essentially the same since it is not a significant classifier.

4.2. Turkish

In Turkish, overall, the Strong and Weak words were correctly distinguished only 61% of the time by the BLRA with all the acoustic measures, thus much less effectively than in French. As Table 2 shows, Duration is indeed not a significant classifier in Turkish, however, Pitch and Intensity are also both quite weak. In fact, with regard to Pitch only mean F0 is significant, not ΔF0. (Centralization classified the targets correctly 54% of the time, but this is not relevant for the proposal investigated here.)

Table 2. *BLRA results for Turkish.*

Overall Classification	**Individual Predictors**	**Classification for Individual Predictors**
61%	▪ Mean F0 ▪ Intensity	58% 57%

The z-scores in Figure 2 provide further insight into the acoustic properties under investigation.

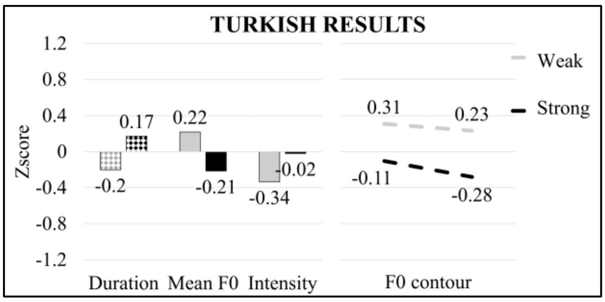

Figure 2: *Turkish - Duration, F0, Intensity and Contour Results. Textured pattern and dashed lines indicate non-significant properties.*

With regard to Duration, as in French, there is an increase in the Strong items, however, it is not significant in Turkish. As in French, the mean F0 is lower in the Strong items, though the (falling) contours are essentially the same; indeed, ΔF0 is not a significant classifier in Turkish. As predicted, and differently from French, Intensity is somewhat higher for the strong items than the weak ones.

5. Discussion

The data from French and Turkish provide a comparison of the acoustic properties of stressed vowels in Strong and Weak positions in prosodic structures. We are thus able to both examine the way stress is manifested in S vs. W positions within each language, and compare the ways the two languages make use of the various properties associated with stress. The findings then permit us to assess the prediction made by the PPH that the way a language expresses stress corresponds to its basic word order, and this in turn, can serve as a bootstrapping cue for infants. Specifically, as previously proposed, it was predicted that our data would reveal duration as the primary cue for stress in French, but pitch in Turkish (possibly with the addition of intensity). That is, the W-S pattern associated with VO word order would be mirrored in an iambic prominence pattern favoring duration in French, while Turkish would not rely on quantity due to the S-W pattern of its OV word order, mirroring trochaic stress.

Indeed, our findings support the prediction that Turkish would not rely on Duration as a stress property. This was seen in the BLRA, where Duration was not a significant classifier. Instead, F0 (58%) and Intensity (57%) were significant, but neither yielded a strong classification (given that chance is 50%) and the overall distinction between the S and W positions was also not very strong (61%). Furthermore, the z-scores show that the F0 distinction is actually due to a pattern that runs contrary to expectations, with F0 being higher in the W than the S context.

Our findings also show that, as expected in an Iambic pattern, Duration does play a role in distinguishing Weak and Strong elements in French. The interpretation is complicated, however, by several additional observations. First, despite the significant finding for Duration, both of the Pitch measures (mean F0 and ΔF0) show substantially stronger classifications (88%, 83%) than Duration (63%). Moreover, as in Turkish, F0 is lower in the S than the W context, whereas the expectation is that prominence would result in enhanced (i.e., raised) F0.

It should be noted that F0 was also reported to be a significant prominence property in French in [1], but this was dismissed on the grounds that the F0 property did not co-occur with Intensity, which is considered to be the main correlate of Trochaic structure. Another consideration may also be relevant, however: although the basic word order of French is Weak-Strong (Head-Complement), some of the targets used in the experiment involved an alternative structure in which the Complement precedes the Head (cf. Ajd-N in (1) above). There may thus have been some intermingling of the acoustic properties in question.

Finally, closer examination of the experimental structure used in [1, 2] shows that it essentially corresponds to our Strong context. That is, all of the stimuli in [1, 2] appeared in a position within a carrier sentence that resulted in the addition of strong (IP) boundary properties to the final word, thus to the S word in French, but to the W word in Turkish. Since our experimental structures explicitly compare focused and non-focused targets, the patterns observed in the Strong context reflect the presence of the IP boundary, while those in the Weak context eliminate it. We now have an explanation for the lower F0 on the S words in both French and Turkish in our experiment, as well as in the S French words in [1, 2]: they are due to an IP-final decline in F0, as seen clearly in the contours in Figures 1 and 2. Thus, instead of dismissing the French F0 finding since it lacks a corresponding Intensity pattern [1], we see that the combination of these two properties actually provides additional insight into the nature of the structures in question. That is, the lack of increased Intensity on the S word is consistent with its position at the end of the IP, rather than a reflection of the use (or not) of Intensity in signaling stress. At the same time, it is also possible that the higher F0 reported for the S words in Turkish in [1], is in fact at least partially due to the IP boundary. That is, rather than the Complement (i.e., S word) being enhanced by raised F0, it may be that the Head (i.e., W word) is lowered because of its IP final position. This would yield an apparent trochaic (S-W) pattern, but it would be due to a phenomenon other than the manifestation of stress. Finally, if there is an IP boundary at the end of the stimuli in [1, 2] we must raise the question of whether the greater Duration associated with the Strong (Complement) position in French might also be due to its IP final position, as opposed to the use of this property for stress. In fact, this was considered in [1] with a comparison of German VO and OV structures; however, since the targets were produced sentence-finally, this did not eliminate the possible confound of a strong boundary.

6. Conclusions

In the present investigation of the Phrasal Prominence Hypothesis, we further tested whether the different acoustic properties of Iambic and Trochaic stress patterns carry over to the W-S and S-W phrasal prominence patterns of VO vs. OV structures in French and Turkish, respectively. If so, these properties could then be taken as cues to basic word order in a language, and thus serve as bootstrapping tool for pre-verbal infants. Instead of finding that Duration was the main prominence property in French and Pitch the main property in Turkish, we found that Pitch properties were the most effective cues in both languages. While duration was also a (weak) significant cue in French, it is possible that that is due to the additional presence of an IP boundary following the target. Indeed, a boundary pattern was also observed in the lowered F0 in both languages, and possibly in the failure to find increased intensity in French since the stressed item could instead exhibit a drop in intensity at the end of the IP.

In sum, we do not find the crucial differences in the manifestation of stress in French and Turkish that would reliably lead to the identification of their basic VO and OV word orders. Nevertheless, our findings for all three properties (i.e., Duration, Pitch, Intensity) are consistent with broader IP boundary marking, and thus may still be viewed as supporting the most basic and general tenet of prosodic bootstrapping - a correlation between basic syntactic structures and prosodic phenomena. As Fernald and McRoberts [3] caution, however, we must still consider whether the acoustic properties are robust enough to actually predict particular syntactic structures, as opposed to just being more probable in relation to those structures.

7. References

[1] M. Nespor, M. Shukla, C. Avesan and R. van de Vijver, "Different phrasal prominence realization in VO and OV languages," *Lingue e Linguaggio*, vol. 7.2, pp. 1-28, 2008.

[2] M. Shukla and M. Nespor, "Rhythmic patterns cue word order," in *The Sound Patterns of Syntax*, Oxford; New York, Oxford University Press, 2010, pp. 174-188.

[3] A. Fernald and G. W. McRoberts, "Prosodic bootstrapping: A critical analysis of the argument and the evidence," in *Signal to Syntax: Bootstrapping from Speech to Syntax in Early Acquisition*, Hillsdale, NJ, Erlbaum Associates, 1995, pp. 365-387.

[4] J. Morgan and K. Demuth, Signal to Syntax: Bootstrapping From Speech To Grammar in Early Acquisition, Mahwah, NJ: Lawrence Erlbaum Associates, 1996.

[5] M. van Heugten, I. Dautriche and A. Christophe, "Phonological and prosodic bootstrapping," in *Encyclopedia of Language Development*, Thousand Oaks, CA, SAGE Publications Inc, 2014, pp. 447-451.

[6] B. Höhle, "Bootstrapping mechanisms in first language acquisition," *Linguistics*, vol. 47, no. 2, pp. 359-382, 2009.

[7] M. Nespor, M. T. Guasti and A. Christophe, "Selecting word order: The rhythmic activation principle," in *Interfaces in phonology*, Berlin, Akademie Verlag, 1996, p. 1–26.

[8] M. T. Guasti, M. Nespor, A. Christophe and B. van Ooyen, "Pre-lexical setting of the head-complement parameter through prosody," in *Approaches to Bootstrapping: Phonological, lexical, syntactic and neurophysiological aspects of early language acquisition, Volume 1*, Amsterdam - Philadelphia, John Benjamins, 2001, pp. 231-248.

[9] T. L. Bolton, "Rhythm," *American Journal of Psychology*, vol. 6, p. 145–238, 1894.

[10] H. Woodrow, "Time perception," in *Handbook of experimental psychology*, New York, NY, Wiley, 1951, p. 1224–1236.

[11] H. Woodrow, "A quantitative study of rhythm," *Archives of Psychology* (New York), vol. 14, pp. 1-66, 1909.

[12] G. Cooper and L. B. Meyer, The rhythmic structure of music, Chicago, IL: University of Chicago Press, 1960.

[13] B. Hayes, Metrical stress theory: Principles and case studies, Chicago: University of Chicago Press, 1995.

[14] A. Prince, "Quantitative consequences of rhythmic organization," in *Papers from the Annual Regional Meeting*, Chicago Linguistic Society, 1990.

[15] J. J. McCarthy and A. Prince, *Prosodic morphology*, Unpublished ms., University of Massachusetts, Amherst & Brandeis University, 1986.