Perception of prosodic correlates of Topic and Focus in Albanian

Enkeleida Kapia
LMU, Munich & ASA, Tirana

Alejna Brugos
BU, Boston & Simmons U., Boston

This study reports preliminary results of a pilot perception experiment on Albanian prosody. Using Rapid Prosody Transcription (RPT) [3, 4], we investigated how native Albanian speakers perceive prosodic prominence for topic and focus in naturally produced speech. Kapia [6] suggested that focus in Albanian comprises two syntactic correlates, i.e. rheme (new information) and kontrast (contrastive new information). Following this, [7] offered the first evidence on prosodic correlates of four information structure constructs in Albanian which speakers systematically marked with differing sentence-level prosodic patterns. This study expands on those results: 1) by using more naturally produced (vs. read) speech and 2) with a perception task.

As part of a larger effort to collect spontaneous Albanian speech, 45 native speakers of Albanian were recruited for a production task. Our analysis for this pilot perception task used productions from only one female speaker. Short (~20-second) sound files were taken from longer interviews in which speakers described 2-picture sequences (from QUIS [10]) with context built such that either the agent or the patient was in focus. Two items from each of 3 conditions were examined, (where T=topic, R=rheme, K=kontrast, A=agent, P=patient): 1) new agent, old patient (RA_TP); 2) old agent, new patient (TA_RP); 3) old agent, contrastive patient (TA_KP). 35 naive listeners participated in the experiment online via Percy [5]. Following RPT methods, a text transcript of each file was displayed, with words separated by spaces, but no punctuation. Participants listened and clicked on words they perceived as highlighted in relation to surrounding words. (See Example 1) For each word, a continuous-valued prosody feature was calculated: the proportion of transcribers who marked the word as prominent (the p-score, between 0 and 1).

A preliminary analysis examined p-scores under the 3 tested conditions, shown in Table 1. Figures 1 & 2 show excerpts from 2 conditions with p-scores. A partial analysis of tonal patterns associated with listener-identified prominences shows mild support for categories proposed by [7] (Table 1). TA nouns in [7] frequently showed low f0 aligned with a strong syllable (analyzed as L*, with or without a following rise), and similarly here the single token of L*+H is in the TA condition (TA_RP, for Item 1 in Table 1). P nouns in [7] for both R and K conditions frequently showed f0 falling from a high preceding the target word (analyzed as H+!H* or !H*), and were most often deaccented or !H* in T conditions. Similarly, in this study, P nouns were realized with f0 patterns compatible with !H* and H+!H* analysis, but neither T token showed a very low p-score, as might be expected from a deaccented word.

Relative p-scores for a given A/P pair did not always meet our expectations, nor were they consistent: i.e., T nouns were not always less prominent than R or K nouns. It should be noted that [7] found substantial variation in prominence realization, and used data from multiple speakers. This current study includes only a few examples from one speaker, so no generalizations can yet be made about f0 characteristics of prominence events.

In addition, agreement on prominence marking was generally high, comparable to RPT results for English [3] vs for Malay [9] (which showed vast variability). Further, examination of tokens with high p-scores showed that they were realized with f0 movements compatible with pitch accents, as suggested in [7], thus supporting the hypothesis that Albanian marks prominence acoustically via pitch accents. We plan to conduct a more detailed analysis of the acoustic cues to prosodic structure (including boundaries), including f0 movement (using PoLaR [1]) and non-f0 cues, including voice quality and duration, for more speakers in the corpus. Since work on any aspect of Albanian prosodic structure remain sparse, this study emerges as a step further towards understanding its intonational phonology.

Example 1: A sample fragment of a transcript annotated for prominence. Participants clicked words heard as prominent to turn them red. In this example, "një karrike" was marked as prominent.

> ndërsa në   të dytën ai po   mban   lart *një karrike*
> while   in   second   he pres.part. hold-pres. up   a chair-acc

Figure 1: An excerpt from the TA_RP condition. Words and p-scores are shown time-aligned to the signal (left) and graphed (right). The Agent target (*grua,* "woman") is Topic, and marked in green. The Patient (*dem,* "bull") is Rheme, and marked in orange. Pitch tracking errors, likely due to amodal voice quality, were manually corrected (blue dashed lines).

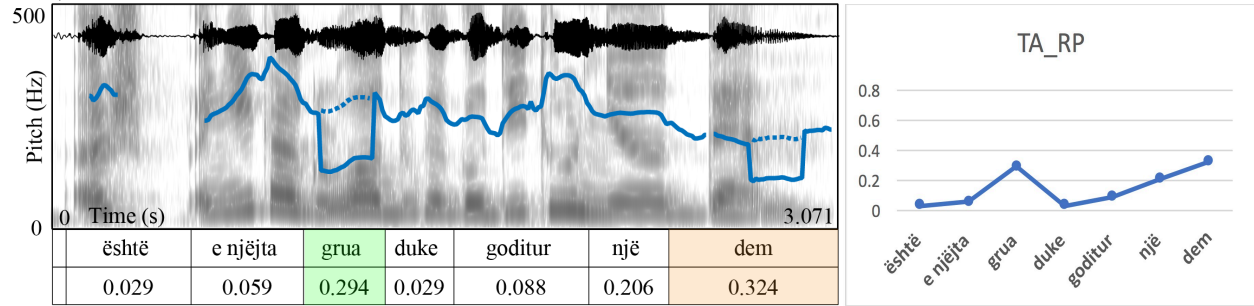| është	e njëjta	grua	duke	goditur	një	dem
0.029	0.059	0.294	0.029	0.088	0.206	0.324

Figure 2: An excerpt from the RA_TP condition. Words and p-scores are time-aligned to the signal (left) and graphed (right). The agent target word (*grua,*) is Rheme, in orange. The patient target word (*demin* "bull", acc.) is Topic, in green.

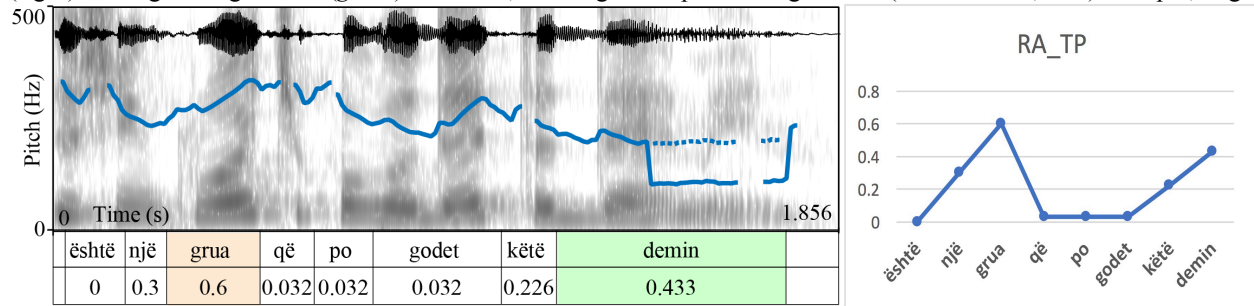| është	një	grua	që	po	godet	këtë	demin
0	0.3	0.6	0.032	0.032	0.032	0.226	0.433

Table 1: Summary of p-score and potential tone patterns following hypothesized pitch accents [7] using using IPrA/ ToBI [8,3] conventions for target Agent and Patient words in the 6 stimuli (3 focus conditions, 2 scenarios/items for each)

	item 1. Agent: *burrë*		item 1. Patient: *karrike*		item 2. Agent: *grua*		item 2. Patient: *dem*	
	p-score	Tone label	p-score	Tone label	p-score	Tone label	p-score	Tone label
TA_RP	1	L*+H	0.129	!H*	0.294	!H* or H+!H*	0.323	H+!H*
RA_TP	0.556	L+H*	0.571	!H*	0.6	L+H*	0.433	H+!H*
TA_KP	0.062	0 (pronoun)	0.483	H+!H*	0.179	H* or L+H*	0.333	H+!H*

References

[1] Ahn, B., Veilleux, N. & Shattuck-Hufnagel, S. (2019). *Annotating Prosody with PoLaR: Conventions for a Decompositional Annotation System*. ICPhS 19. **[2]** Beckman, M., Hirschberg, J. and Shattuck-Hufnagel, S. (2005). "The original ToBI system and the evolution of the ToBI framework," in S.-A. Jun [Ed], Prosodic typology, Oxford, U. Press, 9-54. **[3]** Cole, J., Mahrt, T., Hualde, J. (2014). "Listening for sound, listening for meaning: Task effects on prosodic transcription." *Speech Prosody* 7, 859-863. 2014. **[4]** Cole, J. and Shattuck-Hufnagel, S. (2016). New methods for prosodic transcription: Capturing variability as a source of information. *LabPhon 7*(1). **[5]** Draxler, C. (2011). Percy – an HTML5 Framework for Media Rich Web Experiments on Mobile Devices. *Interspeech 2011*, 3339–40, Florence, Italy. **[6]** Kapia, E. (2010). *The role of syntax and pragmatics in the structure and acquisition of clitic doubling in Albanian*. PhD dissertation, Boston Univ. **[7]** Kapia, E. & Brugos, A. (2016). Information Structure-Prosody Interface: Towards a Model of Albanian Intonational Phonology. *Speech Prosody 2016*, 946-950. **[8]** Prieto, P., Jun, S.-A., & Hualde, J. (2015). IPrA Pitch Accent Labels. Workshop on Developing an International Prosodic Alphabet (IPrA) within the AM framework (Satellite meeting, ICPhS 2015). **[9]** Riesberg, S., Kalbertodt, J., Baumann, S. and Himmelmann, N.P. (2018). On the perception of prosodic prominences and boundaries in Papuan Malay. *Perspectives on information structure in Austronesian languages*, *21*, p.389. **[10]** Skopeteas, S., Fiedler, I., Hellmuth, S., Schwarz, A., Stoel, R., Fanselow, G., & Krifka, M. (2006). *Questionnaire on information structure (QUIS): reference manual* (Vol. 4). Universitätsverlag Potsdam.