

Accented Pronouns and Unusual Antecedents : A Corpus Study

Anubha Kothari

Pitch accents are intonational markers of salience, generally accompanying new and content-worthy information. So what is their function when they appear with pronouns, semantically-impooverished items that stand for old information? Many have suggested that they signal “unusual” resolution, or somehow restrict coreferentiality [1]:¹

- (1) a. John_i called Jim_j a Republican. Then he_i insulted him_j. (*John insulted Jim.*)
b. John_i called Jim_j a Republican. Then HE_j insulted HIM_i. (*Jim insulted John.*)
- (2) a. After he_{i/j} woke up, John_j went to town.
b. After HE_i woke up, John_j went to town.

In particular, [2, 3, 4, 5] have argued that the referent of an accented pronoun is not in the immediate focus of attention, so the accent serves as a cue to its cognitive status. Empirical evaluations [6, 7] have found little support for these claims, although these studies could not investigate the potentially confounding influence of contrast, and suffered from insufficient or inadequate data and statistical analysis. Thus, the question of what pronominal accent is signaling (if anything) and the factors that determine its distribution, remain highly in need of study.

I used a large and richly annotated subsection of the Switchboard corpus of spontaneous conversational speech and analyzed the third-person coreferential pronouns therein via two sets of mixed effects logistic regression models. In the first set of models, I explored the usefulness of various properties of the pronoun and its antecedent in predicting the presence or absence of pitch accent on the pronoun. Amongst the fixed-effects factors were *contrast*, *antecedent-subjecthood*, *antecedent-pronominality*, *clause-distance between antecedent and pronoun*, *pronoun-subjecthood* and *speaker-category* (a rough approximation of three speaker “styles”, namely low-, mid-, and high-accenters). I included two random-effects factors to model the influence of individual differences between *speakers* and between *pronoun-types*. Interestingly, model comparisons and model fit measures indicated that none of the factors that pick out “unusual” antecedents are useful predictors of pronominal accent, whereas *contrast*, *speaker-category* and the *pronoun-type* random-effects factor are extremely significant (all $p < .001$). These results suggest that not only are differences among speakers and pronoun-types key components in explaining the variation in pronominal accentuation, pronominal accent may be signaling contrast rather than something about the attentional status of the pronoun’s referent.

In the second set of models, I limited the dataset to those coreferential pronouns with antecedents in the immediately preceding clause, since most of the attentional proposals discuss pronominal accentuation only in the context of two-utterance windows. Again, antecedent-properties were not useful predictors of accent while *contrast* was. Also, I found no significant relationship between absence of accent and whether the pronoun was a backward-looking center C_b (i.e., referring to the most salient possible entity from the previous clause, approximated as those pronouns which were subjects and had subject antecedents in the previous clause).

The results of this corpus study show that pronominal accent is not a robust cue of an “unusual” antecedent, but may be signaling contrast instead in some cases. Moreover, this study highlights the need for understanding the role of speaker- and item-variation in the accentuation of pronouns, and indicates that it might be useful to discover speaker types or styles. The significance of the speaker- and item- factors illustrates how critical it may be to include them in the analysis of linguistic phenomena more generally.

¹Capitals indicate a pitch accent.

References

- [1] Adrian Akmajian and Ray Jackendoff. Coreferentiality and Stress. *Linguistic Inquiry*, 1(1):124–126, 1970.
- [2] Janet Cahn. The Effect of Pitch Accenting on Pronoun Referent Resolution. In *Proceedings of the Association for Computational Linguistics*, pages 290–293, Cambridge, Massachusetts, 1995.
- [3] Jeanette K. Gundel, Nancy Hedberg, and Ron Zacharski. Cognitive Status and the Form of Referring Expressions in Discourse. *Language*, 69(2):274–307, 1993.
- [4] Megumi Kameyama. Stressed and Unstressed Pronouns: Complementary Preferences. In Peter Bosch and Rob van der Sandt, editors, *Focus: Linguistic, Cognitive, and Computational Perspectives*, pages 306–321. Cambridge University Press, 1999.
- [5] Christine Nakatani. *The Computational Processing of Intonational Prominence: A Functional Prosody Perspective*. PhD thesis, Harvard University, 1997.
- [6] Maria Wolters and David Beaver. What does *he* mean? In *Proceedings of the Annual Meeting of the Cognitive Science Society*, 2001.
- [7] Maria Wolters and Donna K. Byron. Prosody and the Resolution of Pronominal Anaphora. In *Proceedings of the 18th International Conference on Computational Linguistics, COLING2000*, 2000.