PROMINENCE VARIATION BEYOND GIVEN/NEW

Christine H. Nakatani
Bell Laboratories, Lucent Technologies
600 Mountain Avenue, Murray Hill NJ 07974, USA
email: *chn@research.bell-labs.com*

Abstract

Prominence variation is known to be determined in part by discourse factors, such as the givenness or newness of the discourse entity being realized to the discourse. However, few empirical studies have been carried out to explain a wider range of phenomena occurring in natural speech corpora. In this study, corpus linguistics methods are applied to a task-oriented monologue corpus to show that the given/new dichotomy does not explain the data at hand. Rather, there are marked linguistic configurations in which discourse focusing status combines with other linguistic cues to meaning – such as grammatical function and form of referring expression – to convey subtle shades of given/new that can be captured by using a computational discourse model to define attentional focusing status.

Keywords: prosody, prominence, discourse.

1. INTRODUCTION

It is an accepted principle of prosody that new information is likely to bear accentual prominence while given information will not. This corpus-based study reports data from extended spontaneous discourses in which this principle is not directly obeyed. Rather, the majority of expressions, both given and new, are realized with citation-form accentuation. Significant differences in prominence do not arise between the classes of discourse-given and discourse-new items, but do arise when subclasses of given and new items are defined within a computational discourse modeling framework. Further, the marked cases of accentuation show significant interactions with two local focusing features, grammatical function and form of referring expression, extending our present understanding of the roles these features play in determining prominence patterns in discourse.

2. CORPUS AND METHODS

Corpus linguistics methods were applied to nine spontaneous direction-giving monologues from the Boston Directions Corpus [9], produced by a male American English speaker. The corpus consists of spontaneous and read elicited monologues. The speakers provide information for using public transportation to get between points in the Boston area, based on scenarios given by the experimenters. The corpus analyzed for this study totalled 2359 words, comprising 621 referring expressions.

2.1. Coding Referring Expressions

To provide an objective definition of a referring expression, a syntactic definition was used. Minimal, non-recursive noun-phrase (NP) constituents (meaning none of its child constituents are NPs), referred to as baseNPs, were identified using Collin's lexical dependency parser [2]. In the following complex NP, baseNPs appear in square brackets:

> [**the brownstone apartment building**] at [**the corner**] of [**Beacon**] and [**Mass Ave**]

BaseNPs were taken as units of referring expressions and were semi-automatically labeled for grammatical function and form of referring expression.

2.2. Discourse Analysis

BaseNPs also provided the minimal units labeled for coreference relations and attentional status in the discourse. Strict coreference relations amongst the referring expressions were identified by hand using DTT (Discourse Tagging Tool) [1]. The coreference relations were then automatically combined with statistically reliable discourse segmentations [6] to derive the discourse focusing status of referents, using a system defined in [7] based on the focus-stack model of global focusing proposed in [4]. Discourse-new referents not in any coreference relations are assigned to the "single-mention" status. Discourse-new referents that begin reference chains belong to the "first-mention" class. Referents whose closest antecedent occurred in the same, previous, or earlier discourse segment are assigned to the "immediate", "neighboring" and "stack" classes respectively, reflecting the discourse focusing status of the antecedent in the attentional structure. (See also [3]).

2.3. Consituent-based prominence classes

Prominence was analyzed semi-automatically according to a constituent-based scheme [7, 8] that judges the deviation from the citation-form accent pattern for the entire referring expression.[1]

[1] Citation-form accentuation for this study was determined automatically by a TTS system that received each sentence

In this scheme, an expression whose accent pattern matches citation-form, word-for-word, is assigned to the CITATION class. Expressions with fewer or more accented words than citation-form are assigned to the REDUCED and SUPRA classes respectively. If there are mismatches between the citation-form and the actual accenting in both directions, then the expression belongs to the SHIFT class. To summarize, for each baseNP, one of the following four accenting patterns was assigned:

- CITATION FORM: exact match between actual and TTS-assigned word accenting.
- SUPRA: the actual accentuation includes at least one additional accented word than the TTS prediction, and no accented words are predicted unaccented by TTS.
- REDUCED: the actual accentuation includes at least one unaccented word that is accented in the TTS prediction, and no unaccented words are predicted accented by TTS.
- SHIFT: At least one accented word is predicted unaccented by TTS, and at least one unaccented word is predicted accented by TTS.

Table 1 illustrates this accent classification system with several examples from the Boston Directions Corpus.

3. RESULTS

3.1. Distributional Analysis

Distributional analysis of the annotated corpus (see Table 2) showed that the prominence of discourse-new entities (i.e. single-mention and first-mention) did not differ significantly from discourse-old entities (i.e. immediate, neighboring and stack focus) ($p<.32$, $\chi^2=3.52$, df=3). However, statistically significant differences arose between the prominence classifications of the five separate focusing statuses ($p<.03$, $\chi^2=23.95$, df=12). The variation appears to arise not in the distribution of citation-form expressions, which is 70-80% across focus classes, but in the distributions of supra, reduced and shift expressions ($p<.008$, $\chi^2=21.05$, df=8).

When citation-form expressions are factored out, the distributional trends are as follows: (1) single-mention expressions are more likely to be supra than reduced compared to the overall population, while (2) first-mention items are more likely to be reduced; (3) immediate focus items exhibit the most shift prominence patterns, and also more supra prominence than reduced; (4) neighboring focus items display similar distributions of supra and reduced prominence; (5) stack focus items are much more likely to be reduced than supra, compared to the overall population.

3.2. Given Items

Interpreting the results for "given" items (i.e. immediate, neighboring, stack), it appears that the more distant the antecedent in the hierarchical discourse focus structure, the more likely prominence reduction becomes. This is based on the fact that immediate focus items are on the whole more prominent than neighboring focus items, which are in turn on the whole more prominent than stack focus items. This data challenges the claim that items that are re-introduced to a discourse are generally made more prominent than those that remain in focus [5, 11].

3.2.1. Form of Referring Expression

The finding that prominence "fades" as more "given" items are re-introduced is a curious one. However, an explanation lies in the significant interaction of form of referring expression with the deaccentuation of given items in general. Of all reduced items, there is a significant tendency (Table 3, $p<.0002$, $\chi^2=14.85$, df=1) for *given* items to be realized by proper name expressions (91%); in comparison, *new* items are realized by proper name expressions only 52% of the time.

DISCOURSE STATUS	FORM OF REFERRING EXPRESSION	
	Name	Other than name
Given	35	3
New	17	16

Table 3: Interactions of form of referring expression, prominence and discourse focusing status.

The discovery of this interaction is consistent with the theoretical claim made in earlier related work [7] that for reduced, proper name expressions, the form of expression conveys the relative newness of the discourse item into the attentional model, while the lack of prominence conveys the fact that the item has been attended to in the prior discourse. The coupled gradation of givenness and accentual reduction is nonetheless a novel and unpredicted result that extends previous empirically based results.

3.3. New Items

As for "new" items (i.e. single-mention, first-metnion), the given/new principle predicts that first-mention and single-mention items would be similarly prominent. However, while single-mention and first-mention items both occur with citation form prominence about three-quarters of the time, there is a tendency for

in the corpus in isolation. The actual accenting was determined by prosodic labeling using the ToBI standard [10]. Word accenting predictions were produced by the Bell Laboratories NewTTS system [12].

Accent class	TTS-assigned accenting	Actual accenting
citation	*a* LITTLE SHOPPING AREA *we*	*a* LITTLE SHOPPING AREA *we*
supra	*one* *a* PRETTY *nice* AMBIANCE	ONE *a* PRETTY NICE AMBIANCE
reduced	*the* GREEN LINE SUBWAY YET ANOTHER RIGHT TURN	*the* GREEN *Line* SUBWAY *yet* ANOTHER RIGHT TURN
shift	*a* VERY FAST FIVE MINUTE *lunch*	*a* VERY FAST FIVE *minute* LUNCH

Table 1: Examples of citation-based accent classes from the Boston Directions Corpus.

Focus Class	Accent Class			
	Citation	Supra	Shift	Reduced
Single-mention	.73	.16	.06	.10
First-mention	.74	.07	.01	.18
Immediate	.80	.11	.03	.06
Neighboring	.77	.11	–	.12
Stack	.70	.11	.18	.18
Average	.76	.12	.01	.11
Total N	471	73	9	68

Table 2: Distributions of discourse focusing status for each accent class.

single-mention items to be otherwise slightly more prominent than first-mention items, as evidenced by the supra and reduced distributions. Further, immediate focus items are overall more prominent than either of the new item categories. Thus, the simple dichotomy of given/new fails to explain the data at hand.

3.3.1. Grammatical Function

Further distributional analysis reveals that grammatical function plays an important role in determining the prominence of discourse-new items. When the single-mention populations of supra and reduced items are compared, there is a significant tendency for single-mention supra items to appear in adjunct position (83%). Single-mention reduced items, in contrast, appear in adjunct position only 35% of the time. This difference in distributions is significant at $p<.002$, $\chi^2=10.64$, df=1. (See Table 4).

Single-mention	Grammatical Function	
	Adjunct	Non-adjunct
Supra	24	5
Reduced	6	11

Table 4: Interactions of grammatical function, prominence and single-mention discourse givenness.

However, for first-mention items, the opposite tendency occurs at a significant level (Table 5, $p<.02$, $\chi^2=10.37$, df=3). That is, *reduced* first-mention items occur in adjunct or direct object position 94% of the time, while *supra* first-mention items occur in subject position 50% of the time and in adjunct position 50%.

Thus, grammatical function interacts significantly with discourse focusing status to determine prominence markings based not only on discourse structural features, but also on the lexical semantics that underly argument structure, and the sentence semantics that underly sentential structure. Grammatical function was shown in [7] to make independent contributions to accentuation, apart from discourse focusing status.

4. CONCLUSION

This corpus linguistics study continues earlier work [7], by further examining the rich interactions between attentional modeling, coference relations, grammatical function and form of referring expression. While patterns of prominence in spontaneous discourse that run counter to the given/new principle overall are uncovered, more subtle trends in accentuation can be identified when more sophisticated computational discourse modeling is brought to bear. The results extend the approach of [7], in which prominence is seen to play a role in both local and global discourse focusing processes, and at the same time is constrained in interpretation by the linguistic configurations that dynamically unfold throughout the discourse. A view

First-mention	Grammatical Function			
	Adjunct	Subject	Direct Object	Appositive
supra	3	3	0	0
reduced	10	0	5	1

Table 5: Interactions of grammatical function, prominence and first-mention discourse givenness.

of the given/new hypothesis from the perspective of computational discourse modeling can provide explanations that cover the data and are consistent with previous empirically based principles of prominence interpretation.

5. REFERENCES

[1] C. Aone and S. W. Bennett. Evaluating automated and manual acquisition of anaphora resolution strategies. In *Proceedings of the 33rd Annual Meeting*, Boston, 1995. Association for Computational Linguistics.

[2] M. J. Collins. A new statistical parser based on bigram lexical dependencies. In *Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics*, 1996.

[3] B. Grosz and C. Sidner. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3):175–204, 1986.

[4] B. J. Grosz. The representation and use of focus in dialogue understanding. Technical Report 151, SRI International, Menlo Park, CA, 1977.

[5] J. Hirschberg. Pitch accent in context: predicting intonational prominence from text. *Artificial Intelligence*, 63(1-2):305–340, 1993.

[6] J. Hirschberg and C. Nakatani. A prosodic analysis of discourse segments in direction-giving monologues. In *Proceedings of the 34th Annual Meeting of the ACL*, Santa Cruz, 1996. Association for Computational Linguistics.

[7] C. H. Nakatani. *The Computational Processing of Intonational Prominence: A Functional Prosody Perspective*. PhD thesis, Harvard University, Cambridge, MA, May 1997.

[8] C. H. Nakatani. Constituent-based accent prediction. In *Proceedings of the 36th Annual Meeting of the ACL and 17th International Conference on Computational Linguistics*, pages 939–945, Montreal, Quebec, August 1998. Association for Computational Linguistics.

[9] C. H. Nakatani, B. Grosz, and J. Hirschberg. Discourse structure in spoken language: studies on speech corpora. In *Proceedings of the AAAI Spring Symposium on Empirical Methods in Discourse Interpretation and Generation*, Palo Alto, CA, March 1995. American Association for Artificial Intelligence.

[10] J. Pitrelli, M. Beckman, and J. Hirschberg. Evaluation of prosodic transcription labeling reliability in the ToBI framework. In *Proceedings of the 3rd International Conference on Spoken Language Processing*, volume 2, pages 123–126, Yokohama, Japan, 1994.

[11] A. Sanford, K. Moar, and S. Garrod. Proper names as controllers of discourse focus. *Language and Speech*, 31:43–56, 1988.

[12] R. W. Sproat and J. P. Olive. A modular architecture for multi-lingual text-to-speech. In J. P. H. van Santen, R. Sproat, J. Olive, and J. Hirschberg, editors, *Progress in Speech Synthesis*. Springer Verlag, New York, 1997.