Prosodic Correlates Of Referent Status

PROSODIC CORRELATES OF REFERENT STATUS

Maria Wolters
Institut für Kommunikationsforschung und Phonetik, University of Bonn, Germany

ABSTRACT

Prosodic correlates of 6 referent status taxonomies and 3 distance-from-last-mention heuristics both on the acoustic and on the symbolic level (ToBI) were investigated in a corpus of short news reports read by 6 professional newsreaders. Symbolic correlates are found mainly for pronouns, acoustic correlates for nouns and proper names. However, both form and extent of these correlates varies considerably between speakers.

1 INTRODUCTION

1.1 What is given?

Prosody can have many functions, among these signalling "new" information. It is frequently assumed that discourse-new information tends to be accented, and discourse-old information to be deaccented. But what exactly is this givenness? A popular operationalization is that entities which have already been mentioned in the discourse are given, all others new [11, 5]. Frequently, an item is also regarded as new if the distance to the last mention is more than one sentence. However, givenness of information and givenness of discourse referents are two different concepts [9]. The first one is closely related to focus/background structure, while the second one can be characterized along the two dimensions discourse-old/new vs. hearer-old/new [14] or via accessibility hierarchies such as [3, 7]. To avoid confusion, I will call this second concept "referent status". Its exact formalization depends largely on the specific semantic (or psycholinguistic) theory of dialogue processing assumed. In this paper, I will not commit myself to any specific theory, but instead work with several taxonomies and heuristics which are described briefly in sec. 2.2.

1.2 Previous Work

Many researchers have examined prosodic correlates of referent status. Brown [2] concluded from a sample of short task-oriented dialogues that a simple given/new dichotomy was a good predictor for accent position, while more detailed taxonomies such as [14] have syntactic correlates. Nooteboom and Kruyt [11] found that it is more acceptable to accent expressions referring to given entities than to accent those referring to new ones. Accents on given referents can even provide valuable cues for discourse management. For example, Nakatani [10] concluded from an analysis of monologue that accents on pronouns can signal center shifts. Duration is another potential correlate of referent status. Second occurrences of words in communicative discourse, be it read or spontaneous, tend to be shorter than first occurrences [6]. In her study of short read texts from a single speaker, Eefting [5] found no direct effect of givenness on word duration, but an interaction between givenness and accentuation. Lastly, if there are clear prosodic correlates of givenness, we should not expect them to be the same across speakers [18], which raises the question if some speakers use prosody less optimally than others. Experimental testing of such claims requires clear hypotheses about the role of prosody in anaphora resolution and referent processing, which is beyond the . . .

1.3 Scope of this paper

In this paper, I examine how the speakers of the Boston University Radio News Corpus [12] use prosody to signal referent status, examining both symbolic correlates (pitch accent location and type) and acoustic correlates (duration and fundamental frequency) with respect to 9 operationalizations of referent status. The study focuses on 3 questions: are there any correlates of referent status in the corpus, if so, which taxonomies have the strongest correlates, and how large is inter-speaker variation?

2 DATA

2.1 The corpus

The corpus consists of the four labnews texts from the BU Radio News Corpus which were read by 6 FM radio newsreaders (3 female, f1a,f2b,f3a; 3 male, m1b,m2b,m3b). f2b, m1b, m2b, and m3b usually edit the texts they read. The total text length is around 2100 words. Almost all of these texts have been labelled with ToBI (for an introduction, see [1]), with the exception of f1a and f3a, where prosodic labels for a few paragraphs were not included in the version of the corpus I used. In ToBI, words are annotated with pitch accents, phrase tones, boundary tones, and break indices. This study focuses on the type and presence of pitch accents, which are written as combinations of high (H) and low (L) tones. Other information used were part-of-speech (POS) tags (Penn Treebank 1 tagset, [12]), syllable prominence (binary), text transcriptions and F0 curves (computed by Entropic ESPS Waves).

The Boston corpus was chosen for several reasons: Firstly, the texts are long enough to contain chains of more than two coreferring noun phrases, i.e. NPs which share the same referent. Secondly, the structure allows inter-speaker comparisons. Furthermore, the corpus has been used frequently for prosody research (cf. eg. references in [16]).

Status	Definition
brand-new unanchored	new to hearer and discourse
brand-new anchored	new, but anchored in old referent
unused	header-old, discourse-new
situationally accessible	referent present in situation
textually accessible	referent mentioned before
inferentially accessible	referent related to old ref. sub-/superset, part/whole, frame
active	ref. currently talked about

Table 1: Taxonomy S

Lastly, this corpus is a real challenge for theories relating accents to givenness, because such correlations have been notoriously hard to determine, at least for f2b, the best studied speaker [16, 8]. Thus, I expect that weak correlations in this data may surface more strongly in other types of read speech and maybe even more in spontaneous speech.

Read speech was chosen because it allows to study coherent texts with well-defined sentence and paragraph boundaries, structural information which is valuable for prosody generation. (For this corpus, the texts were transcribed manually from recordings [12].)

2.2 Annotation of the Texts

First, coreference chains were annotated according to the Message Understanding Conference Coreference Task Definition [4] standard. All coreferring NPs were then annotated with

a) four classes of syntactic information: function (subject, object, adjunct); modifiers (adjective, relative clause, prepositional phrase); head noun; presence and type of article

b) two taxonomies of referent status: S, derived from [14, 9], (Tab. 1), and GH (Givenness Hierarchy [7], Tab. 2).

The most problematic of these attributes is clearly referent status. Poesio and Vieira [13] report that even for the relatively straightforward and well-documented annotation schemes they experimented with, in their case restricted to definite descriptions, inter-annotator agreement is rather low. Should therefore all research on empirical correlates of information status rest until theories can be developed that allow consistent annotation schemes? Obviously not, because such results are important for language understanding and generation. One way out is to restrict oneself to simple operational definitions of two to three levels of referent status. Alternatively, a small team of labelers can try to arrive at a consensus labeling of a corpus based on clear guidelines and frequent discussion.

These problems were addressed as follows. First, since it was not possible to have several annotators label the complete data, all annotations were inserted manually by the author in four separate steps, adding one layer of annotation at a time. Secondly, four coarser taxonomies were derived from S: KS (brand-new, accessible, active, unused),

Status	expression specifies
type identifiable	class of referents
referential	unique referent intended
uniquely identifiable	unique referent identifiable
familiar	known referent
activated	ref. in working memory
in focus	ref. in focus of attention

Table 2: The Givennness Hierarchy

SS ([17], old, mediated, brand-new), D (discourse old/new, [15]) and H (hearer old/new [15]). These classifications are progressively less fine-grained and therefore less susceptible to labelling errors. Finally, three distance heuristics were computed based on the annotated coreference chains: distance from last mention in paragraphs (DP), distance from last mention in sentences (DS), and the ternary KD (first mention, last mention in current or previous sentence, last mention two or more sentences ago).

3 METHOD

The bulk of the analysis is restricted to nouns and pronouns in referring noun phrases, because our definition of givenness really only applies to them. Acoustic analyses and inter-speaker comparisons were only conducted for the four fully annotated speakers.

3.1 Target variables

I examined both potential symbolic and acoustic correlates of the 9 taxonomies and heuristics described in Sec. 2.2. The symbolic target variables are derived from the ToBI annotation. In the analyses, the accent classes were both preserved as labelled (classes: H*, !H*, L*, H+!H*, L*+H,L+!H*,L+H*,*, with ! signalling downstep) and reduced to 'low', 'downstepped', 'high', 'unknown' according to the class of the starred tone. The category unknown was added because it is quite frequent for m1b. I will report mainly results on accent type and placement.

The acoustic variables cover the prosodic parameters pitch and duration. Two duration variables were measured, total word duration divided by the number of phonemes and normalized to z-scores (zdur), and duration of the longest syllable nucleus (maxdur). F0 is covered by three variables: F0-maximum of the nucleus of the pitch accented syllable, (accF0) F0 range of that nucleus (measured as distance between F0 minimum and maximum, accrg), and the highest F0 on a syllable nucleus (maxF0). Additionally, the quality of the longest syllable nucleus, the number of prominent syllable nuclei, and the total number of syllable nuclei were extracted.

3.2 Control variables

Major phrase boundary effects are covered by a position variable which can take on any of the 5 values 'paragraph-initial', 'sentence-initial', 'paragraph-final', 'sentence-final', and 'medial'. Other control variables are

speaker, POS, function in coreferring NP (head, head of a referring modifier, other), and syntactic function. Semantic aspects are rather crudely covered by the POS-level distinction between nouns and proper names. Contrastivity was not controlled for, because it is notoriously difficult to define properly.

3.3 Statistical method

All statistical analyses were conducted with the package R (http://www.cis.tuwien.ac.et/R). The influence of each taxonomy and heuristic was tested separately. For acoustic target variables, I used general linear model with position, syntactic function, function in coreferring NP, and accentuation as independent variables and an additional term for the interaction between accentuation, position and information status. Accentuation is a binary variable which states whether a word carries a pitch accent or not. For symbolic targets, a series of Fisher (pitch accent location) and χ^2 (pitch accent type) tests were performed for each speaker and word class (noun, proper noun, pronoun). In the noun tests, I also varied syntactic function and function in coreferring NP (head of NP or a modifier, modifier).

4 RESULTS

4.1 General observations

Three of the four control variables strongly influence accentuation. The odd one out is syntactic function, for effects are rather low (e.g. pronouns (PP) tend to be accented in object position, $p < 0.1$, Fisher test). Few pronouns bear a pitch accent (19,4%), while 85,54% of all proper names (NP) and 79,12% of all other nouns (NN) are accented. For preposed genitives, the difference is even larger: 87,75% accents on proper names, and 58,7% accents on nouns. Because of these effects, nouns, proper names, and pronouns are analysed separately. Words in heads of referring NPs are also accented rather frequently (71,8%).

In fact, speakers vary greatly both in the percentage of accented words and in the type of accents they use, confirming the results of a previous comparison of 3 speakers reading the same story [16]. While the correlation r between speakers f2b, m1b, and m2b for accent position is around 0.6/0.7, it never exceeds 0.3 for m3b. Both m1b and m2b accent over 57% of all words in the text. But frequency alone cannot explain why m1b accents only 12,5% of all pronominal subjects, while m2b accents 29,3%. For pitch accent type, the situation is not much better (cf. Table 3). m3b uses disproportionally many H+!H* accents, while f2b and f1a have the highest percentage of L*. The most frequent pitch accent is generally H*, except for m1b, where it is * (no labelled tone).

4.2 Symbolic correlates of referent status

If we look at the effect of the taxonomies and heuristics on pitch accent type and placement for both nouns and pronouns, almost all have significant effects. But a closer analysis reveals that most of this is due simply to the deaccentuation of pronouns. In fact, pronouns are the only class which, when analysed separately, show effects of referent status for four speakers (Tab. 4).

	*	L*	H+H!	L+!H*	!H*	% acc.
f1a	4.8	**6.1**	2.3	1.1	**16**	48.7
f2b	3.2	**6.7**	0.6	3.1	**17.1**	**50.1**
f3a	6	2.5	3.5	5	11.6	55.65
m1b	**70.8**	2.7	1.5	0.08	**3.2**	57.25
m2b	4.1	3.2	5.5	**5.6**	8.6	58.67
m3b	6	0.6	**10.8**	**5.9**	10.75	**50.64**

Table 3: Accent frequency relative to total no. of accents and frequency of accents relative to total no. of words

Speaker	f2b	m1b	m2b	m3b
Tax.	GH,KS,S	S,SS	D,KD,S	SS
% acc. PP	14	14,14	28,3	20,41

Table 4: Effects of referent status on pronoun accentuation at $p < 0.05$ or better; no such effects for f1a and f3a.

Two classes of pronouns are almost never accented: pronouns followed by a clitic verb and "you", which is mainly used for general statements. There are no pronouns which are accented most of the time, except for "one" in the context "one of them". But this use is more cardinal than pronominal. Cataphoric pronouns also tend to be accented, especially by m2b. Non-referential pronouns are accented less frequently than referential ones (10% vs. 20% in the complete corpus).

Distance metrics fare worse than the more elaborate hierarchies GH and S. In fact, a series of two-sided t-tests shows that the distance of accented discourse-old nouns, proper names, and pronouns to the last mention does not differ significantly from that of unaccented ones.

We might also expect that heads of NPs introducing referents which are mentioned more than once tend to be accented. The only relevant effect in the data points in the opposite direction: the shorter a reference chain, the less likely it is that an NP head of the first mention bears an accent ($p < 0.05$, two-sided t-test). Furthermore, head nouns of subsequent mentions are more likely to be accented if they belongs to a short coreference chain. A possible explanation is that as frequently mentioned referents become more salient, they receive less pitch accents.

For pitch accent type, the situation is even more complex. For proper nouns, only f2b, f3a, and m2b show any significant effects; for nouns, only f1b, f3a, and m2b show any effects and these mostly for KD. There is no one accent that signals givenness; rather, accent types tend to be more or less frequent depending on referent status. Tab. 5 shows the effect of D ($p < 0.05$ for all three speakers) on the distribu-

	!H* D-new	!H* D-old	L* D-new	L* D-old
f2b	**28.9**	9.4	2.4	**12.5**
f3a	15.5	**31.2**	0	0
m2b	8.8	**16.9**	3.3	2.8

Table 5: Distribution of L* and !H* in terms of total accented D-new/D-old proper names.

	maxdur	zdur	accrg	accF0	maxF0
f2b	3	3	3	3	1
m1b	1,3	1,2		1,2,3	2
m2b	1	1	2,3	2,3	2
m3b	1,2,3	1,2,3	1,2,3	1,2,3	

Table 6: Acoustic correlates for POS classes. Key: 1 – NN, 2 – NP, 3 – PP

tion of L* and !H*. While the magnitude of the percentages reflect general accentuation patterns, we note that f3a and m2b tend to use !H* on D-old referents, whereas f2b shows exactly the opposite pattern.

4.3 Acoustic correlates of referent status

The only consistent pattern is that there are far more significant acoustic than symbolic correlates of referent status. Tab. 6 summarizes which variables correlate at $p < 0.05$ with any of the taxonomies from Sec. 2.2. Results are given for each POS class and speaker.

Acoustic correlates tend to be strongest for the two durational variables (several taxonomies show significant effects at $p < 0.05$), with the exception of m2b, where intonation dominates. Referent status influences pronoun accentuation; this explains the rather consistent correlates which accrg and accF0 show for pronouns. All male speakers show a wide range of acoustic correlates of referent status for nouns and proper names. An inspection of the full results suggests that these are stronger for nouns than for proper names. m2b mainly uses duration for nouns, and F0 for pronouns and proper names. He also tends to make D-new nouns shorter than D-old ones. One reason could be that those noun phrases which introduce discourse referents contain more modifiers [19] which can help in referent identification.

5 DISCUSSION

In our corpus, prosodic correlates of referent status are both difficult to find, and vary greatly from speaker to speaker. There are three possible reasons for this. First, linguistic correlates may be so conclusive that prosody is not really needed for successful referent resolution. This is highly probable for these texts [19]. Secondly, prosodic correlates of referent status should also depend on speaking style; for example, we would expect them to be more important in conversation. Thirdly, the observed acoustic effects for nouns and proper names could be mainly due to a reduction of articulatory effort for easily retrievable items [6], whereas pragmatic factors come into play when accenting pronouns.

ACKNOWLEDGMENTS

Thanks to Thomas Portele, Marc Swerts, and Jacques Terken for comments during earlier stages of this work, Bertjan Busser for help with the data and ILK Tilburg (especially Sabine Buchholz) for their hospitality.

REFERENCES

[1] M. Beckman and G. Ayers. Guidelines for ToBI labeling, v. 2.0, 1994.

[2] G. Brown. Prosodic structure and the Given/New distinction. In A. Cutler and D.R. Ladd, editors, *Prosody: Models and MEasurements*, 1983.

[3] W. Chafe. Cognitive constraints on information flow. In R. Tomlin, editor, *Coherence and Grounding in Discourse*, pages 21–51, Amsterdam etc., 1987.

[4] Message Understanding Conference. The MUC coreference task definition v. 3.0. In *Proc. of the 6th Message Understanding Conference*, 1996.

[5] W. Eefting. The effect of "information value" and "accentuation" on the duration of dutch words, syllables. *JASA*, 89:412–424, 1991.

[6] C. Fowler and J. Housum. Talkers' signaling of "new" and "old" words oin speech and listeners' perception and use of the distinction. *J. Memory Language*, 26:489–504, 1987.

[7] J. Gundel, N. Hedberg, and R. Zacharski. Cognitive states and the form of referring expressions in discourse. *Language*, 69:274–308, 1993.

[8] J. Hirschberg. Pitch accent in context: Predicting prominence from text. *Artificial Intelligence*, 63:305–340, 1993.

[9] K. Lambrecht. *Information Structure and Sentence Form*. CUP, Cambridge, 1994.

[10] C. Nakatani. Constituent-based accent prediction. In *Proc. ACL/Coling Montreal*, pages 939–945, 1998.

[11] S. Nooteboom and J. Kruyt. Accents, focus distribution, and the perceived distribution of given and new information: an experiment. *JASA*, 82:1512–1524, 1987.

[12] M. Ostendorf, P. Price, and S. Shattuck-Hufnagel. The Boston University radio news corpus. Technical report, Boston University, 1995.

[13] M. Poesio and R. Vieira. A corpus–based investigation of definite description use. *Computational Linguistics*, 24:183–216, 1998.

[14] E. Prince. Toward a taxonomy of Given/New information. In R. Cole, editor, *Radical Pragmatics*, New York, 1981. Academic Press.

[15] E. Prince. The ZPG-letter: Subjects, definiteness, and information–status. In W. Mann and S. Thompson, editors, *Discourse Description*, pages 295–325, Amsterdam, 1992. Benjamins.

[16] K. Ross and M. Ostendorf. Prediction of abstract prosodic labels for speech synthesis. *Computer Speech and Language*, 10:155–185, 1996.

[17] M. Strube. Never look back: An alternative to centering. In *Proc. ACL/COLING Montreal*, pages 1251–1257, 1998.

[18] M. van Donzel and F. Koopmans-van Beinum. Pitch accents, boundary tones, and information status in spontaneous discourse in dutch. In *Proc. ESCA Workshop on Intonation*, pages 313–316, 1997.

[19] M. Wolters. Linguistic correlates of referent status. Technical report, IKP, Univ. Bonn, ftp://asl1.ikp.uni-bonn.de/pub/mwo/lingkorr.ps, 1999.