PROSODY-PARTICLE PAIRS AS DISCOURSE CONTROL SIGNS

Dafydd Gibbon and Claudia Sassen
Fakultät für Linguistik und Literaturwissenschaft
Universität Bielefeld, Postfach 100151, D-33501 Bielefeld
Tel. +49 521 106 3510, FAX: +49 521 106 6008
E-mail: gibbon@spectrum.uni-bielefeld.de, csassen@hrz.uni-bielefeld.de

ABSTRACT

We address the problem of integrating the description of discourse particles and their intonation into an HPSG–based lexicon for spontaneous speech applications, and propose a lexical sign type called *prosody–particle pair* which has similar structure to grammatical inflexions and is formally described as a nested attribute–value structure. In our discussion we generalise a known class of 'stylised intonations' to include the level intonation of hesitation particles.

Previous descriptions of discourse particles, their roles and their relations to intonation, have been informal. Our proposal is the first to model particle–intonation relations explicitly and in detail as an inflexion–like complex sign in a formal lexicon. The inclusion of the intonation of hesitation phenomena in the class of stylised intonations on formal and functional grounds is also new.

1. DISCOURSE CONTROL SIGNS

Prosody and discourse particles jointly function as complex discourse control signs with a variety of functions. Our use of the term 'sign' is related to that of HPSG [10]. A sign is a four-dimensional structure <*SYN*, *DTRS*, *SURF*, *SEM*> with two *compositional* dimensions (syntactic distribution in the immediate linguistic context, and internal components), and two *interpretative* dimensions (surface, consisting of phonetic and orthographic interpretation, and semantic, consisting of truth-functional and pragmatic contextual substructure).

We concentrate on signs with the *phatic* function ([6], [12]) of establishing, maintaining and terminating a *communication channel* (*CC*). We introduce *CC* as a substructure of *SEM*, with further substructure <*MEDIUM*, *CONTACT*, *DIRECTION*>. For example, *MEDIUM* is further specified for [STORAGE={transient,permanent}] and other properties, *CONTACT* is specified *inter alia* for <*FUNCTION*, *ROLE*, *CONTROL*>, *ROLE* has substructure specifying social and interactional discourse roles including *client, chairman, addressee*; other speaker and hearer attributes may be needed for specific purposes (cf. Chapters 3 and 8 in Gibbon & al. [5]).

The state of *CC* changes before, after and within dialogue turns, depending on *Dialogue Syntax* (*DS*) conditions, such as speech act cycles with 'back–channel handshaking' (e.g. *uptake–securing, repair, acknowledgement, appraisal*). We represent these structures as nested attribute-value matrices.

We develop a description of *prosody–particle pairs* (*PPPs*) which is integrated into a prosodic extension [1] of attribute–based inheritance lexicon theory of the Generative Lexicon type [11].

2. STYLISED INTONATION

Stylised intonation has been described in detail in many studies on intonation (cf. especially [3] and [7]). The auditory impression is of chanting or singing; stylised contours are closely related to the melody of simple children's chants such as English 'Cowardy cowardy custard, your face is made of mustard!', or German 'Angsthase, Pfeffernase, morgen kommt der Osterhase'. The contour typically occurs in the context of discourse *openings* and *closings* in public contexts such as calls ('Coo–ee!', 'Mo–ther!'), greetings ('Hello–oo') and farewells ('Bye–ee!'). There are language specific variants of form and function. A rising variant of form occurs in English calls, but not in many other languages. In German, the contour is also used in certain repair contexts ('Lau–ter!', '"Jo–hann" habe ich gesagt!'); in French, stylised contours appear to be used in certain speaking styles in other contexts, but we are not aware of research on this.

Apart from the possible case of French, the functions of stylised intonations can all be classified as *phatic*, i.e. *CC*–oriented functions, in the sense introduced above. Stylised intonation is easy to classify phonetically and linguistically, as it differs considerably from the modulated and level pitch contours of continuous speech in the following ways:

1. Tonal *stability* as (a) single monotone, or (b) monotone pair. The main subtype of (b) has falling pitch with an interval ratio (Hz) of about 6:5, i.e. a minor third or quarter octave [9], with the perceptual impression of 'chanting' or 'singing', and possibly with a behavioural correlate in *auditory feedback control* (cf. [4]).

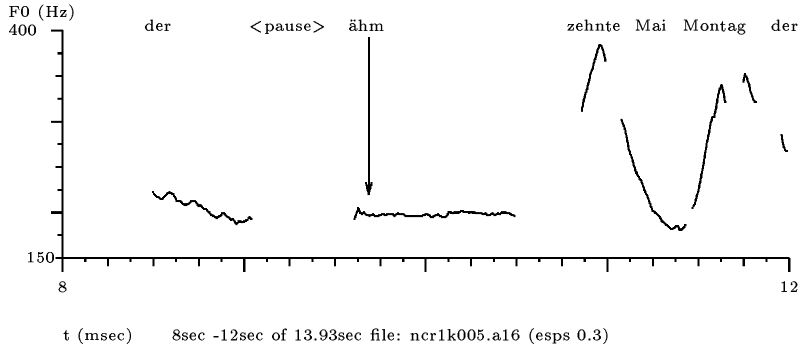

Figure 1: German discourse particle 'ähm' with stylised F_0 trajectory.

2. Distributional *independence* from sentence structure, though cooccurring with it.
3. A number of *phatic* functions including *CC*–based discourse *CONTROL* functions in dialogue openings and closings.

We claim on the basis both of our *CC* analysis and on acoustic grounds that when level pitch, acoustically a flat F_0 trajectory, is paired with discourse–medial hesitation particles, this is also a case of stylised intonation. An example[1] of a particle–prosody–pair of this kind is shown in Figure 1.

3. PARTICLE AND PROSODY POLYSEMY

Any given discourse particle such as English *erm*, German *ähm* is multiply polysemous and occurs in a variety of CC-dominated contexts. The most obvious example is *yes*, which can either be specified for an INDEX attribute, and interpreted as assigning the same truth value as the preceding speaker to a propositional schema in the preceding turn, or simply specified for a CONTROL attribute, as confirmation of having heard (but not necessarily having understood or being in agreement with) the preceding turn.

The same applies to intonation patterns. A rising nuclear tone can (but need not) signify a question, a non–final element of a list, a subject constituent, simply that a turn is to be continued, or a polite (perhaps subservient) termination to a turn with a subsequent turn expected. A common feature of *discontinuity* may be shared by these functions, or a looser functional family resemblance may be postulated.

Combinations of particles and prosodic patterns are, however, vastly less polysemous; particles constitute discourse control signs, for instance, only when paired with specific intonations in context. It may be argued that the polysemy in prosody–particle pairs is analogous to morphological syncretism, with lexicalised *PPP*s consisting of a particle complement of a 'prosodic inflexion' superfix head; in the examples discussed here, the prosodic superfix has the *turn–medial* phatic function *channel–sustaining*. More generally, we suggest that utterances of all types only occur paired with prosodic contours, and that our approach can be extended to signs in general.

The superfixation operation is analogous to grammatical inflexion by affixation, in which an inflexional head with contextual, often deictic relevance, is attached to a stem. Composition is by the prosodic relation of temporal *overlap* ($X \circ Y$), not the *precedence* ($X \prec Y$) relation which would interpret a concatenation operation (see [2]). There is an analogy to the spatial overlap operation with typographic highlights in orthography. Intonation is treated as *head* on account of its greater generality and its contribution to the interpretation of the whole sign.

4. PROSODY–PARTICLE PAIR SIGNS

In order to describe prosody–particle pairs as signs, we represent the four–dimensional structure of signs, in general following HPSG conventions, as in Figure 2. We diverge from the HPSG conventions in that we add additional prosodic substructure to the PHON attribute, and adopt attributes from Generative Lexicon Theory (cf. [11]) for the SEM attribute.

Figure 2 shows a composite lexical entry for an abstract lemma HES (hesitation) as a nested attribute–value structure. The composite entry has two parts, joined by the compositional operator '$\otimes$': first, an item of type *prosody*, second an item of type *particle*. In the context of the type *prosody* under the SURF attribute, the operator is interpreted as temporal overlap, '$\circ$'. Under the SEM attribute, the operator is interpreted as unification, '$\sqcup$'.

In the attribute–value model, we abstract the compositionality constraint for the type *prosody–particle–pair* out of the lexical item. In HPSG terms, the con-

[1] VERBMOBIL scheduling dialogue corpus CD-ROM 1, dialogue n005k, female speaker CR1, turn 005.

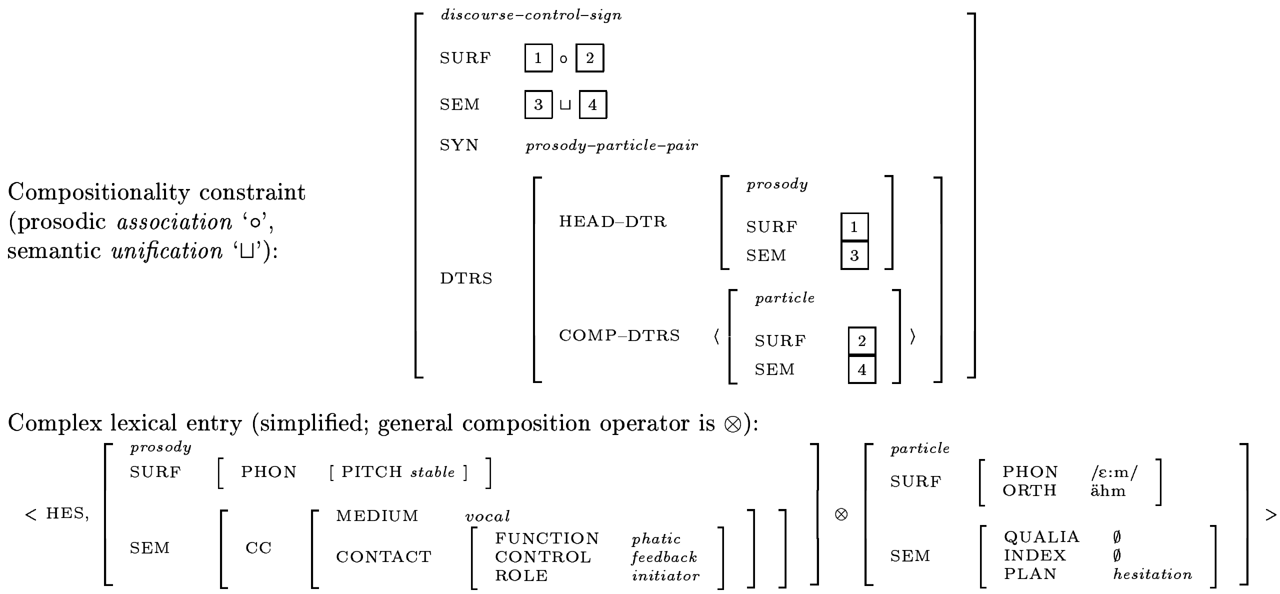

Figure 2: Prosody–particle pairs as lexical types.

straint has the status of a general 'principle of grammar'. Attributes which are not required in specific lexical entries are omitted, yielding an underspecified attribute–value structure which will unify with other attribute–value structures, but with fewer constraints than a fully specified structure.

5. SIGN ATTRIBUTES

Pragmatic attributes such as CC and its substructure are included under the attribute SEM. The attributes PHON, and also in principle ORTH, are included under SURF.

The type *prosody* has the pragmatic attribute CC, with its substructure CONTACT, in turn subdividing into the attributes FUNCTION, whose values are ultimately defined under speech act and discourse theory, reduced here simply to *phatic*, CONTROL, with functions in turn–taking, framing (dialogue, episode and turn structuring) and speech style (formality), here reduced to the *feedback* function, and ROLE, i.e. the relation of participants to each other at various levels, here reduced to *initiator*.

For the type *prosody*, the SURF attribute is not subdivided for ORTH, as there is no correlate of stylised intonation in written language. The attribute SURF only has PHON substructure. For some prosodic forms it can be argued that there is an orthographic interpretation in terms of punctuation, highlight fonts, and layout; this is not at issue here. As space does not permit a presentation of the entire 'feature geometry' of the PHON attribute, the main relevant substructure of PHON for this example is taken to be PITCH, with the value *stable*; this must be enhanced by other features in a full description.

The type *particle* is more like a conventional word, for which SURF also has a substructure ORTH. In some particles, such as *uhuh*, *mhm*, PHON structure does not correspond to standard phonotactics, and ORTH represents a partially conventionalised onomatopoeic spelling.

For the type *particle*, the SEM attribute has the features which a lexical item would be expected to have. To indicate this, the attributes QUALIA and INDEX are included, even though the values for this particular entry are empty (and thus, strictly speaking, the attributes could have been omitted). The QUALIA (content) attribute is relevant for non–hesitation particles with turn–introducing function such as *right*, *sure*, *yes*, *no*. It may be argued that the particles with Boolean function such as *yes*, *no* and related items, may be specified for the INDEX attribute, since they anaphorically assign a truth value to a propositional schema in the preceding turn.

6. PHATIC MONITORING SIGNS

In his three–level theory of speech production ('Conceptualiser' / 'Monitor', 'Formulator', 'Articulator', the last two being paralleled on the perceptual side by the 'Speech Comprehension System' and the 'Audition' component), Levelt [8] ascribes a function of *monitoring* in the form of *perceptual loops* to speech production. Levelt proposes two loops, an *internal loop* from the output of the Formulator (the internal phonetic plan or the phonetic string) to the Conceptualiser via the Speech Comprehension System, and an *external loop* from the output of the Articulator (overt speech) to the Conceptualiser via the Audition component and the Speech Comprehension System.

We suggest including a low–level feedback mechanism at the Audition level in the model (or a corresponding loop in a connectionist model) to account for the auditory feedback which permits the production of stable stylised intonations, chanting and singing. This monitoring can then be interpreted as a procedural 'icon' for the more general kind of phatic monitoring in which these contours are involved. In Levelt's terms (p. 497):

> Speakers attend to what they are saying and how they say it. They can monitor almost any aspect of their speech, ranging from content to syntax to the choice of words to the properties of phonological form and articulation.

Interestingly, Levelt discusses intonation in the context of discourse repairs and 'editing expressions' (repair marking discourse particles) (pp. 495ff.), but does not mention flat, stable, stylised pitch contours.

But stylised intonations are not the only prosodic types with phatic function. A very frequent device is *iteration*, as in the following examples:[2] *No, no, no, there's people that handle both* ..., or ... *putting on a shirt or uh, uh, uh, reaching the back seat* ...

In other cases, the first syllable (or syllable fragment) of a non–particle (or, if a monosyllable, the whole word) may be iterated, as in ... *the research in- instead of* ..., *Right I I I think there's been a mistake* ... or *uh so I don't know I don't know do you have* ...

The prosodic property of *iteration* has apparently not been discussed before in the literature. It immediately raises several new questions which it will not be possible to deal with in detail in the present context:

1. How can *iteration* be modelled as a property of discourse repair signs? Perhaps compositionally, as a regular expression with a variable to be filled by the segmental content? This solution harmonises well with recent developments in finite state phonology.
2. Why does *iteration* apply more often to non–particles (e.g. monosyllabic function words or function word sequences, or to initial syllables) than to particles?
3. Why do these iterated units typically occur with stable, stylised pitch contours? How do these relate to lengthening?
4. Should *iteration* be treated as a lexical property, or as a property of signs in general, or is it rather a procedural property of the speech production process (a kind of 'handshaking protocol' for the monitoring subprocess)?
5. If *iteration* is a procedural property of the monitoring subprocess, does the same also apply to stylised intonations and lengthening?

[2] Taken from a telephone dialogue corpus, cf. [12].

7. CONCLUSION

We have developed a description of a 'phatic' lexical sign type, which we call *prosody–particle pair*, with similar structure to morphological inflexions. These signs are formally described as nested feature structures, and are embedded in a hierarchy of *prosody*, *particle* and *prosody–particle–pair* types. We have also generalised a known class of 'stylised intonations' to include pitch on hesitation particles, and identified a new prosodic *iteration* function as a phatic marker.

REFERENCES

[1] Bleiching, D. (1994). Integration von Morphophonologie und Prosodie in ein hierarchisches Lexikon. In: Trost, H., ed., *Proc. KONVENS '94, Verarbeitung natürlicher Sprache, Wien, September 1994.* Berlin: Springer. 32–41.

[2] Carson-Berndsen, J. (1993). *Time Map Phonology and the Projection Problem in Spoken Language Recognition.* PhD thesis, University of Bielefeld, April 1993.

[3] Gibbon, D. (1976). *Perspectives of Intonation Analysis.* Bern: Lang.

[4] Gibbon, D. (1988). Intonation and discourse. In J. Petöfi, ed., *Text and Discourse Constitution.* Berlin: de Gruyter. 3–25.

[5] Gibbon, D. & al. (1997). *Handbook of Standards and Resources for Spoken Language Systems.* Berlin: Mouton de Gruyter.

[6] Jakobson, R. (1960). Closing statement: Linguistics and Poetics. In T. A. Sebeok, ed., *Style in Language.* Cambridge, Mass., MIT Press. 350–377.

[7] Ladd, D. R. (1978). Stylized intonation. In *Language*, 54. 517–540.

[8] Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation.* Cambridge, Mass.: MIT.

[9] Liberman, M. Y. (1975). *The Intonation System of English.* PhD thesis, MIT, Cambridge MA.

[10] Pollard, C. & I. Sag (1987). *Information–based Syntax and Semantics.* Stanford: CSLI.

[11] Pustejovsky, J. (1995). *The Generative Lexicon.* Cambridge, Mass.: MIT Press.

[12] Sassen, Claudia (1996). *Phatische Kommunion: Begriffsbestimmung anhand englischsprachiger Telefonate und eines Rundfunkinterviews.* M.A. Thesis, U Bielefeld.

[13] Stenström, Anna–Brita (1994). *An Introduction to Spoken Interaction.* London, New York: Longman.