

Book review

From Data to Speech. Language Generation in Context

Mariët Theune: Eindhoven: Eindhoven University of Technology, 2000. Thesis, x + 209pp.

Theune's book is devoted to a very large set of issues from different layers, ranging from natural language processing to linguistic theory. The author's main aim is to elucidate ways in which context can and should be taken into account when constructing procedures of text generation. This goal, as well as the degree of attention paid in the book to the exploration of theoretical issues concerning the anchoring of language in the process of communication, issues which have become urgent with the 'interactive turn' in linguistics, make the book highly interesting.

The book is divided into six chapters, in the first of which (Introduction, p. 1) an important requirement is already formulated, namely that spoken text should be generated (from non-linguistic information in a database), rather than just its written counterpart. This makes it necessary to work with enriched text, which has to contain markers indicating "the desired prosody of the spoken output" (cf. also p. 18 and Sections 2.5 and 3.6). It may be added that this (still often overlooked) goal of connecting research in grammar with that in sentence prosody, is crucial, since (i) the written form of text is just a simplified, incomplete way of rendering speech (the prototypical manifestation of language), and (ii) when studying sentences as anchored in context, it is crucial to reflect those relevant features of prosody that often are not present in writing. Among the further tasks, the author stresses the handling of referring expressions both in what concerns co-reference and the choice of the expression as specifying certain properties of the referent (p. 6f).

Chapter 2 is devoted to a description of the Data-to-Speech (D2S) system the author applies; the author presents a clear characterization and appropriate illustrations of this system on pp. 19ff (it may be added that the restricted number of syntactic templates used is made possible by the schematic nature of the contents of the texts generated). D2S is a generalization of van Deemter and Odijk's Dial Your Disc system (an interface to a music database) and is used by Theune for two applications of different kinds, with output in Dutch. The first of them generates spoken reports on football matches from tabular data, while the other (discussed in detail in Chapter 5) generates dialogues within a public transport information system. While the language generation module of D2S produces 'enriched text' in the sense

mentioned above, the speech generation module turns this representation into a speech signal. The division of these two modules constitutes a substantial innovation in the architecture of text-generation systems. Since the assignment of sentence prosody within the language generation module requires not only ingredients such as syntactic templates, but also input data relevant for the distribution of accents, a procedure determining whether a lexical occurrence carries new or given information is presented (pp. 40ff; see also Sections 5.3.1 and 5.3.3 on two degrees of givenness). If Steedman's remark that "prosodic structure does not always adhere to the traditional subject-predicate division of a sentence" is quoted in a footnote (fn. 5, p. 40), it should have been added that such lack of adherence concerns, as is well known, also other constituent boundaries. The question remains, why theories reducing the weight of constituent structure (that elaborated by Steedman himself, as well as those based on syntactic dependency) are left without discussion, although some of them have brought results that are useful for an analysis of the impact of context on sentence structure and on co-reference, as the author herself states in Chapter 4.

The account of contrastive accent (including its automatic detection), which is presented in Chapter 3, is very useful in that it investigates one of the crucial features of sentence structure which have been paid little attention up to now. The author's summaries of and comments on approaches such as those by Prevost (contrast as based on choice among alternatives), Pulman (relationships to parallelism) or van Deemter (contrariety) are important, and so is her experiment (described on pp. 71ff) with subjects indicating which versions of text segments they found most natural-sounding. The experiment yields convincing results, according to which the presence of a contrastive accent prevails in the case of an item used in parallel with another one in two consecutive sentences, although it does not prevail in the cases lacking this parallelism.

Unfortunately, Theune is not much interested in how contrast and similar phenomena have been analyzed from phonetic and acoustic viewpoints, i.e. having their means of expression in view; it is infelicitous that she works only with a single kind of accent, assuming that there is "no phonological difference between accents that express newness and those that express contrast" (p. 73). In some connections this single accent is called 'contrastive accent' (esp. in Chapter 3), in others, the accent concerns the focus of the sentence. In Theune's framework, these two functions are more or less identified. Her claim that "accent is assigned only to phrases which are marked as being in focus," is conditioned by the fact that in her terminology "a phrase is regarded as being in focus when at least one of its words expresses new or contrastive information" (p. 53, cf. also p. 40). Even if we leave aside the not quite exact use of 'phrase' (this term might be understood as covering also the whole sentence), it can be seen that this makes it difficult—without a distinction between different kinds of accent—to account for the functional difference between *she* and *him* in such well known examples (*He called her a Republican.*) *Then she insulted HIM.* The present reviewer prefers to distinguish between a rising contrastive accent, marked here by (the absence of) italics, and a primarily falling accent characteristic of focus, marked by capitals; although being aware of several kinds of

non-prototypical phenomena due to which ‘rising’ and ‘falling’ can be understood here only as simplified denominations, the reviewer does not want to go into a further discussion of these phenomena. The contrastive *she* would rather be understood, in different approaches to the topic-focus articulation (or information structure) to belong to topic, cf. other examples, such as *This I gave to MOTHER* (that, *to FATHER*), or *Him she SAW*, in which the pronoun starting the sentence hardly could be viewed as belonging to its focus. A criterion useful in this context may be seen in the effect of sentential negation: in *This I didn’t give to MOTHER* the pronoun *this* clearly is not “under negation”, i.e. is not in the focus, and this is similar e.g. with *him* in *Him she didn’t SEE* (us *she didn’t LISTEN TO*).

Theune’s discussion on the presence of syntactic or semantic parallelism belongs to the parts of her book in which she presents new insights, deserving further discussion, although not all her formulations probably can be accepted without hesitation. Thus, e.g. in her ex. (29) on p. 81, ... *I had not expected to see him again and HE looked pretty surprised at seeing ME* (in which again, the rising contrastive stress on *he* might be distinguished from the falling focus stress on *me*) at least one layer of parallelism in the cognitive content can be acknowledged: the core of the events expressed by either of the two sentences is referred to by the verb *see* (with different modalities), and the two participants of the event switch their positions, which may be seen as a legitimate background for contrasting them (i.e. not only that one chosen as the focus, but also that included in the topic). Thus, this example is not altogether excluded from sentence pairs expressing contrastive data structures as specified by the author’s definition of contrast (p. 83). This definition, although not fully adequate (as the author notes on p. 92), constitutes one of the significant contributions to the discussions on contrast and on the impact of context.

Discussing the generation of descriptions in context, in Chapter 4 the author presents her characterization of and comments on Dale and Reiter’s Incremental Algorithm as an efficient way of describing how definite descriptions are specified in accordance with the Gricean maxims, on the basis of a hierarchy of preference of attributes used in a given domain. To distinguish an entity newly introduced to the discourse, it is proposed to use salience weights, first just with the specification of the most salient entity displaying a given property (p. 106f). Then, 10 degrees of salience are distinguished on the basis of a definition derived from Hajičová (whose 1993 book, published in Prague by Charles University, is mistakenly characterized as a journal article in the bibliography), and this approach is compared with that of centering theory. Unfortunately, this comparison is carried out without a deeper discussion, with an illustration by a single example. An experiment (described by Verhagen) is then characterized, in which a set of hypotheses on the preference of different kinds of anaphoric descriptions is scrutinized. This results in confirming the preference of (i) anaphoric descriptions that contain fewer properties than their antecedents and (ii) pronouns, if they can refer to only one entity. A resulting modification of the generation algorithm is briefly presented, with due attention given to rules concerning contrastive accentuation.

Language generation in a spoken dialogue system is discussed in detail in Chapter 5, with an illustration of a procedure including dialogue management and speech

generation as applied to dialogues concerning a public transport information system.

In conclusion, we may state that Theune's book will be welcome as a highly stimulating contribution to discussions concerning issues in which contextual factors are of impact for methods of natural language processing. The author presents very clear, though critical characterizations and well illustrating examples of several existing systems from the domain of text generation, some of which she only comments upon, while others have been chosen as starting points of the author's own approach. Her observations on such problems as the significance of sentence prosody for semantics (and thus also for text generation) are extremely important in this context, since up to now these problems have been neglected in many theoretical approaches, as well as in the domain of applications.

On the other hand, it should also be noted as a drawback of the book that the author's theoretical background is somewhat restricted, since she pays little attention to linguistic discussions that have been devoted to information structure and are important for issues she deals with. This concerns concepts that already have been elaborated in individual theories of the relevant aspects of sentence structure and discourse patterns, be it in the approaches based on classical European structural linguistics (Daneš, Firbas, Halliday), on its formalized continuations (Hajičová), or on post-Chomskyan frameworks such as those by Selkirk, Steedman, Valduví and others. Thus, the notion of contextual boundness (as a reflection of 'givenness' in language structure, which does not always immediately correlate with the presystemic understanding of the term), would be useful in discussing examples such as *My son wants a DOG, but I am ALLERGIC to dogs* (p. 10, with the reviewer's way of indicating the different kinds of accent), in which the deaccentuation of the last word is not due directly to the fact that the concept 'dog' was mentioned before, but rather to the contextual boundness of its second occurrence [which would be contextually nonbound, if connected with a choice among alternatives, e.g. in *My son wants a DOG or a CAT, and I brought him a DOG (since I am allergic to cats)*].

Even so, the book under review certainly will play an important role in further discussions on ways in which the impact of context is to be reflected when constructing procedures of text generation. There is no doubt that this domain of research belongs to the most urgent ones at present, when the interactive nature of language and thus the anchoring of sentences in context, i.e. in the process of communication, has become the center of attention of both theoretical linguistics and its computational applications.

Petr Sgall

Prague, Czech Republic

E-mail address: sgall@ufal.mff.cuni.cz

Petr Sgall (born 1926) studied Linguistics at Charles University, Prague; degrees: PhDr in 1949; CSc (Ph.D.) in 1955; docent (assoc. prof.) in 1958; DrSc in 1966; professor in 1990; now: Professor Emeritus. Awards: A.v. Humboldt research award (1993), Honorary Doctorate of the Institut National des Langues

et Civilisations Orientales, Paris (1995), Honorary Doctorate of the Hamburg University (1998) Main areas of research and teaching: syntax and its relationships to semantics and pragmatics, typology of languages, computational linguistics, sociolinguistics. Main publications (books): *Infinitive im Rigveda*, Prague, 1958; *Generative description of language and the Czech Declension* (in Czech) Prague, 1967; with co-authors: *A functional approach to syntax*, New York: American Elsevier, 1969; *Topic, focus and generative semantics*, Kronberg: Scriptor, 1973; *The meaning of the sentence in its semantic and pragmatic aspects*, Dordrecht: Reidel and Prague: Academia 1986; *Variation in language: Czech as a challenge to sociolinguistics*, Amsterdam: Benjamins 1992; *Topic-focus articulation, tripartite structures, and semantic content*. Dordrecht: Kluwer 1998.

PII: S0378-2166(02)00101-7