textual economy in generation

the SPUD system

Magdalena Wolska

assumption: generation is *goal-oriented*

an agent has certain *communicative intentions* it aims to fulfill (e.g. identifying an entity, communicating a property of an entity)

explicit means of goal satisfation: assemble appropriate syntactic constituents (e.g. entity: use np, action: use vb)

textual economy: implicit goal satisfaction – exploit receiver's recognition of inferential links to material elsewhere in sentence to fulfill independent goals
⇒ *overloading*

GRE: generating a description that characterizes the target entity uniquely

- input: an entity as specified by KB of information that is presumed to be part of *shared knowledge* (common ground)
- output:
  semantics of a linguistic description based on which the receiver of the information will be able to unambiguously distinguish the entity from other salient candidates (distractors)
  surface form that matches the conceptual material

„economy” in GRE : generate a description *D* to indentify *r* in context set *C* just in case KB **supports** attribution of *D* to *r* and **motivates** *D* in describing *r* in *C*

support: there must exist information that r fits the description (else description may not be understood or generated)

motivation: there must be a reason to use each of the properties to describe *r* (else the generator may leave the desriptor out)

cf. *full-brevity* GRE: use as few properties as possible; when multiple descriptions apply, f-b: produce a more complex description when shorter don't refer uniquely

r1 : the big black shaggy dog
r1 : the affable purebred racing dog
the black racing dog

indredients of a GRE component:

- context set of entities that are explicit (or implicit) in discourse
- formalisation of properties of these entities
- specification of how to combine the properties into linguisitc description
- representation for processing
- reasoning to online assess how receiver will interpret the given description
  ⇒ resource intensive

typical GRE scenario:

- any fact to be communicated must fit into abstract grammatical structure (including lexical items)
- any reference to domain entity must be elaborated into a description that uniquely identifies (cf. incremental descrption)
- surface form must be found
  ⇒ step-wise process
  (assuming e in common ground, 1/ identify set of concepts that together distinguish e from its distractors, 2/ derive syntactic structure that realises these concepts)

SPUD: perform syntactic and semantic processes *simultaneously*

References

Matthew Stone and Bonnie Webber. Textual Economy through Close Coupling of Syntax and Semantics. Proceedings of INLG 1998, p. 178--187.

Martha E. Pollack. Overloading intentions for efficient practical reasoning. Nous 25(4):513--536.

see also other references at: http://www.cs.rutgers.edu/~mdstone/nlg.html

you can download SPUD at: http://www.cs.rutgers.edu/~mdstone/class/taglet/
(for instructions on running SPUD, you may contact Magdalena Wolska magda@coli.uni-sb.de)