Lexical Choice and Knowledge Representation

Stephan Busemann

Deutsches Forschungszentrum für Künstliche Intelligenz
Project DISCO
Stuhlsatzenhausweg 3
D-6600 Saarbrücken 11
busemann@dfki.uni-sb.de

March 12, 1992

1 Introduction

Recently the problem of choosing communicatively adequate lexemes has attracted much interest in the NL generation community, in general, the task amounts to deciding for a given representation of an intended meaning, which words will most appropriately convey that meaning to the addressee.

Whether lexical choice must be exact in the sense that all and only the intended meaning is verbalized, depends on the respective communication situation. In a multimodal discourse, where language is supplemented by gestures or graphics, the linguistic device need not convey everything to the partner. In written discourse without a predefined context, as in DISCO, exact verbalization seems much more in order.

In all theories of lexical choice, the *convergence problem* has to be solved: there is always a decision for *exactly one* lexical item.

We may distinguish the following subtasks of lexical choice:

Definite reference, proforms: Events and objects must often be described using words that allow for an unambiguous identification of the referent. The problem subdivides in finding appropriate words for the referents and in describing the relations between them, as deictic and intrinsic readings of "The ball is infront of the car" suggest.

Social jugdement: Some words carry social jugdements with them. German *Putzfrau* and *Raumpflegerin* mean both *cleaning woman,* but only the latter is now used officially.[1] The former has a pejorative connotation. See [6].

Collocations: There are different kinds of cooccurrence restrictions between lexemes. Some words cannot be used together with others, some tend to be used together with others and some yield a different meaning when used with certain others (idioms).

[1] *Raumpflegerin* reminds at *Krankenpflegerin (nurse).*

Choice of open class words: Given a conceptual representation of the intended meaning, an appropriate word for each concept must be identified.

In this paper, we assume that for lexical selection the following kinds of knowledge are necessary:

- the concepts of the meaning representation language
- lexical entries (lemmata and/or phrasal items including semantic and syntactic in formation, among other things)
- knowledge about the reader (including the reader's goals ans beliefs)
- knowledge about the linguistic, situational, and social context

We will show that lexical choice requires a domain model based on linguistic considerations, and that standard KL-ONE techniques are insufficient for parts of the task at hand.

2 Techniques for the Choice of Open-Class Words

We briefly sketch the techniques employed for the choice of open-class words, as described in [13]. In many generation systems, conceptual knowledge is used: however, the criteria for modelling this knowledge largely differ. From a theoretical point of view, one can represent conceptual knowledge as a theorie of mental categories. In implemented generation systems, however, a model is usually oriented towards the special purpose of the respective system. Obviously the concrete task of choosing open-class items depends on the structure of the underlying knowledge base.[2]

Direct replacement: This still often used technique presupposes a one-to-one relationship between concepts and lexical items. For a given concept (e.g. TRUCK) the lexeme *truck* is uniquely determined. This approach does not really deal with lexical choice.

Structural replacement: Partial structures of the meaning representation are identified that match with lexical entities, and the former are replaced by the latter. The procedure terminates if all elements of the meaning representation are replaced by lexical material (cf. [10]).

Classification: The unique lexeme is searched that expresses closest the meaning of the concept to be verbalized. A well-known early approach used decision trees [4] where possible verablizations of a concept (e.g. *eat, drink, breathe* for INGEST) are represented as leaf nodes. More recent work uses classification in KL-ONE based systems.

[2] Experiences with the modelling task for the LILOG-System are described in [8].

Structural replacement and classification can be combined by using classification techniques during pattern matching. The disadvantage of the techniques presented is that they are restricted to considering the propositional content of the meaning representation. It is difficult to account for the assumed knowledge of the reader, the goals of the author, contextual knowledge and the maxims of conversation [5]. Hence, the choice of the most specific lexeme may lead to a correct but inadequate response (1b2) if *bachelor* is more specific than *man*.

(1) a. Is Kim a woman? bl.
No, Kim is a man. b2. No,
Kim is a bachelor.

3 Demands placed on the model

3.1 Noun choice

Definite reference and pronominalization requires taxonomic reasoning, as the following examples show.

(2) We'll take the big truck and the Mercedes. Both vehicles are available.

(3) The big truck is smashed. The engine doesn't work.

(4) Grind the carots and potatos, cook them to a paste and put them into a prewarmed bowl.

(2) involves generalization (choosing a word whose meaning subsumes those of *truck* and *Mercedes).* It is neither the most specific one (e.g. MOTOR-VEHICLE) nor a very general one (e.g. THING). Rather, it is the most specific *basic-level class* [14]. To solve this selection problem by classification, a domain model should include a class of basic-level objects.

In (3) the part-of relation holding between TRUCK and ENGINE allows for a definite description. The representation of the parts of a typical truck is required, but not of additional parts of some special truck in order to avoid the generation of sentences like (5).

(5) ? The small truck of Mr Evens is smashed. The rudder came off.

Example (4) involves pronominalization even though the referents undergo a change of state, as is determined by the verbs *grind* and *cook to a paste.* Obviously a changed aggregate state must not be referred to by the verb, as the following examples suggest.

(6) Melt the ice and put it into a bowl.

(7) * Melt the ice and pour it into a bowl.

The interrelation between event representation and changes of state has been neglected so far, but see [15].

3.2 Adjective choice

Consider the generation of dimensional adjectives such as *high, long, deep*. On the basis of the defining features of e.g. a pole alone it is not possible to verbalize its maximal axis. Whether (8) or (9) is appropriate depends on the pole's position given by the context; only if the pole is upright, (8) is possible.

(8) The pole is 10 m high.

(9) The pole is 10 m long.

[9] suggest a two-staged propositional semantic representation that relates language-independant conceptual entities to lexical entities. Among other things, conceptual features of spatial objects (POLE has a maximal axis) are combined with lexical constraints of dimensional adjectives *(high* requires, in contrast to *long,* a vertical orientation of an object's maximal axis).

The selection of dimensional adjectives also depends on the speaker's spatial location. [9] claim that their theory purports to any dimensional expression, thus including prepositions, nouns, and verbs as well.

For this approach to be used in a lexical choice system, the model of spatial knowledge must be geared towards distinguishing the two levels carefully.

4 Inalienable Possessives

Inalienable possessives allow in German definite descriptions without explicit prior introduction of the referent (10).[3] While in (11) the owner of the ladder is verbalized as a genitive attribute, in (10) it is expressed as a dative NP. Exchanging the syntactic construction would lead to inacceptability in both cases.

(10) Hans trat Martha auf den Fuss. [Hans stepped Peter on the foot.]

(11) Hans trat auf Marthas Leiter. [Hans stepped on Martha's ladder.]

There are, however, no clear boundaries between what must go with a dative or with a genitive. (12) seems to be acceptable with both constructions.

(12) Hans trat Martha gegen das Auto. [Hans stepped Martha against the car.]

A domain model should exhibit the information which objects count as inalienable possessives.

[3]The brackets contain an interlinear translation.

5 Demands placed on the model and the formalism

5.1 Verb choice

The event of buying a car (meaning represented as (13)) can be described by lexically converse verbs (14-17). In these examples, all participants in the buying event (buyer, seller, goods, money) can be verbalized, but not all are obligatory. If a participant is not verbalized, its existence in the event being described can be deduced. As [3] notes, the verbs bring different participants into *perspective*. Verb choice must thus take into consideration which discourse referents are in perspective. This is determined by dynamic context knowledge. [7] demonstrates how lexical and conceptual knowledge represented in one and the same formalism (Ace, extending KODIAK [16]) can be interrelated for verb choice.

(13) CommercialExchange(goods: carl23, buyer: Peter,
seller: John, money: $800)

(14) John bought a car from Peter for $800.

(15) Peter sold a car to John for $800.

(16) John paid $800 to Peter for a car.

(17) Peter received $800 for a car from John.

The domain model must hence be capable of expressing entities in perspective as well as different *views* of some event [7]. A view is a verbalization under certain conditions. For instance, if the seller and the goods are in perspective, the view of (13) as a *selling* is most appropriate. Views relate lexical items to concepts. There is no obvious way how view can be represented in standard KL-ONE dialects [2].

5.2 Noun choice revisited

If the knowledge of dialogue partners is to be considered during lexical choice, we have a domain model[4] and various user-dependant models of how concepts may be verbalized. The system FN [13] helps a user with the decision whether a given object should be used in an action. Should, for instance, a certain flight be used during a journey? FN wants to express that flight ABC3465 lands in La Guardia. It can do this by explicitly saying (18). Alternatively FN could say (19) because FN knows that shuttles typically land in La Guardia. It will depend on the user's knowledge about shuttle flights whether FN will choose the (preferred) shorter version.

(18) Take flight ABC3465 at 11, which lands in La Guardia.

(19) Take the 11 o'clock shuttle.

Since dialogue participants may have different knowledge about shuttles, a lexical choice system should be capable of anticipating (and avoiding) false communicative implicatures

[4] Let us assume that this model be shared by all dialogue participants.

[5] (e.g. by choosing *shuttle* if the partner thinks there will a meal be served, as during other flights, but it won't).

These and related issues are implemented using an "overlay" for the domain model that contains all the user-dependent knowledge [13, 12]. This overlay can be exchanged if a different partner is addressed. Implementing such an overlay will extend the standard KL-ONE formalism.

6 Conclusion

By looking at lexical choice in generation we have demonstrated that a number of demands should be placed on a domain model by developers of NL front ends. Application-oriented models usually do not exhibit sufficient knowledge to fulfill these demands. A step in the right direction (but not yet a solution to the problems mentionned) would be the use of an Upper Model [1], which structures the most general part of the knowledge base according to linguistic criteria.

Moreover we have shown how conceptual and lexical knowledge may depend on each other. Two ways of relating these kinds of knowledge can be pursued. One of them keeps the lexicon separate from conceptual knowledge. This requires the definition of complex interfaces and reasoning processes (a choice may fail due to constraints in either the lexicon or the conceptual knowledge). Views would be part of the interface. The alternative is to explore work in lexical semantics (e.g. [11]) which suggests to incorporate taxonomic distinctions into the lexicon. This requires, in practice, the reconstruction of much of what has been done in KL-ONE style knowledge representation work. A concept of views would then involve an extension to the formalism.

It is a yet open question which of the information needed for lexical choice should be provided by the lexicon and which by the domain model.

References

[1] John A. Bateman. Upper modeling: organizing knowledge for natural language processing. In *Proceedings of the 5th International Workshop on Language Generation,* Pittsburgh, PA., 1990.

[2] R. J. Brachman and J. G. Schmolze. An overview of the kl-one knowledge representation system. *Cognitive Science,* 9:171-216, 1985.

[3] C. J. Fillmore. The case for case reopened. In P. Cole and J. M. Sadock, editors, *Syntax and Semantics Vol. 8. Grammatical Relations.* Academic Press, New York, 1977.

[4] N. M. Goldman. Conceptual generation. In R. C. Schank, editor, *Conceptual Information Processing.* North Holland, Amsterdam, 1975.

[5] H.P. Grice. Logic and conversation. In P. Cole and J.L. Morgan, editors, *Speech Acts,* pages 41-58. London, 1975.

[6] E. H. Hovy. Unresolved issues in paragraph planning. In R. Dale, C. Mellish, and M. Zock, editors, *Current Research in Natural Language Generation,* chapter 2, pages 17-46. Academic, London, 1990.

[7] P.S. Jacobs. King: a knowledge-intensive natural language generator. In G. Kempen, editor, *Natural Language Generation,* pages 219-230. Martinus Nijhoff Publishers, Dordrecht, Boston, Lancaster, 1987. NATO ASI Series E Nr. 135.

[8] Gudrun Klose, Ewald Lang, and Thomas Pirlein. Die ontologie und axiomatik der wissensbasis von leu/2. Iwbs report 171, IBM Deutschland, Stuttgart, Mai 1991.

[9] E. Lang, K.-U. Carstensen, and Geoff Simmons. Modelling spatial knowledge on a linguistic basis. IWBS Report 147, IBM Deutschland, Institut fur wissensbasierte Systeme, Postfach 80 08 80, 7000 Stuttgart 80, December 1990.

[10] J. Nogier and Michael Zock. Lexical choice as pattern matching. In T. Nagle, J. Nagle, L. Gerholz, and P. Elklund, editors, *Current Directions in Conceptual Structures Research*. Springer, New York, 1991.

[11] James Pustejovsky. The generative lexicon. *Computational Linguistics,* 17(4):409-4413, 1992.

[12] Ehud Reiter and Chris Mellish. Using classification to generate text. In *Proc. Conf. of the 30th Annual Meeting of the ACL,* Newark, 1992.

[13] Ehud B. Reiter. Generating descriptions that exploit a user's domain knowledge. In Robert Dale, Chris Mellish, and Michael Zock, editors, *Current Research in Natural Language Generation,* pages 257-286. Academic Press, 1990.

[14] Eleanor Rosch. Principles of categorization. In E. Rosch and B. Lloyd, editors, *Cognition and Categorization,* pages 27-48. Lawrence Erlbaum, Hillsdale, 1978.

[15] Bonnie L. Webber. Accommodating context change. In *Proc. Conf. of the 30th Annual Meeting of the ACL,* Newark, 1992.

[16] R. Wilensky. Kodiak-a knowledge representation language. In *Proc. 6th Annual Conf. of the Congitive Science Society,* Boulder, Colorado, 1984.