Generalized Quantification and Associations

Henk Verkuyl

1. Associations

When I think of Frans Zwarts, I cannot evade thinking about the very long post-strike in November 1983 which prevented Frans' promised comment on a paper from arriving during this strike (and thereafter) and about the beep of a fax awakening me in the midst of the night at about 2.30 a.m. somewhere in 1992 and starting the production of the promised sheets of his part of a joint paper. After which the fax stopped in the middle of page 2. Without a second beep at all, let alone the remaining text.

The inevitability of the connection between the two events and Frans may be taken as one of the elements of what is generally considered an association. It is clear that the two associations just sketched involve memory in a crucial way and that the two recollections push aside other dear recollections that I have of Frans such as those connected with a pleasant stay in Santa Monica, or nearer, with the period that he was a student in Amsterdam, or with my visits to the Petrus Campersingel, etc., etc. The two of the first paragraph are more prominent in my memory and that is why they are probably called associations.

Memory is clearly also involved in the enforced association by Pavlov's dog between the sound of a bell and the expected presence of food, as it was in the case of Proust's famous association between a madeleine and his experiences as a child spending the summer in his parent's country house in Illiers-Combray. In all cases so far, associations do not have much to do with lexical meaning: it is my thinking of Frans, the sound of a bell and the taste of a madeleine which trigger the association. Thus meaning seems to be quite far away from associations, even though the OED illustrates its definition of the word association as 'a mental connection between things' with the example *the word bureaucracy has unpleasant associations*. This raises the question of what linguists might have to do with the notion of association. In other words, does the binary associative relation A(x,y) fall in the domain of linguistic research because x may have words as one of its values?

Looking in the neighborhood, psychology and computational linguistics seem to be the domain in which the nature of associations may reveal itself. Psychologists generally tend to consider associations as connections of some kind between concepts, which are

then taken as the building blocks of our memory. If associations are treated in this way, it means for the relation A(X,Y) that if the word x triggers one or more associations, x is simply to be taken as a label for the concept X. In this way, linguistic semantics is justified to leave the study of associations of the sort described so far completely to the domain of psychology. On that line, it seems wise to stick to the (still sloppy) linguistic notion of connotation and let social psychologists or psycholinguists deal with associations.

I am phrasing this at daggers drawn because in many cases the literature on associations in psychology seems to be unaware of the existence of linguistic semantics. I have here in mind the social psychological research on associations, prejudices and stereotyping done by A.P. Greenwald and many associates.[1] They have developed the so-called Implicit Association Test (IAT) in order to measure the strength of automatic associations between concepts, which are quite generally taken as mental representations of objects in memory. Actually, what they appear to do is to equate the word x and the concept X in the relation A(X,Y) by simply treating words as concepts. In a lot of their experiments the participants are asked to relate pairs of words in order to establish their associative strength on the basis of reaction time. But the words themselves are discussed without any principled distinction between language and domains of interpretation. This makes it possible for formal semanticists to enter the stage again because one could, in principle, use the findings of Greenwald c.s. by defining the standard interpretation function I and, given A(X,Y), by taking I(X) and I(Y) as values in domains of interpretation which are made part of our memory. In other words, the Fregean notion of intension can now be seen as operating on concepts, because it would cover someone's experience with the world out there (stored in concepts) and because it allows for quantification over different domains of interpretation as experienced by an individual. Such a migration of formal semantics to psychology would deprive linguistics of semantics.

There is another field in which the strength of the relation between two (or more) words is measured: computational linguistics as focused on linguistic corpora. In the past decennia, a corpus-based approach to measuring the associative force f (A) in A(X,Y) has made a significant contribution to shifting the original attention of computational linguists from lexicography to theoretical linguistic problems in morphology, syntax, semantics and pragmatics. In the relatively short history of this field, there are two significant stages. One started out with trying to establish relations between words in a text on the basis of co-occurrence and soon one had to distinguish between paradigmatic associations (roughly, associations involving lexical semantic relations such as synonymy, antonymy, hyponymy, etc.) and syntagmatic associations (roughly, verb-argument relations, verb-

[1] See for example, Greenwald et al. (1998), (2002) and Greenwald and Sriram (2010).

preposition relations, etc., but also semantically relevant relations such as between fish and Friday, madeleines and holidays).[2]

A newer approach—maintaining the divide between paradigmatic and syntagmatic associations—focusses on the distance between words in a text in order to measure the strength of A, offering more room for the measurement of the syntagmatic relations.[3] And according to those computational linguists who argue for distance as the criterion for an associative relation, syntagmatic relations are exactly the place to look for if one wants to account for the arbitrariness and the volatility of associations. Which seem to do justice to the associations mentioned at the beginning of the present section.

It is about time to see what formal semantics has to say about associations. In the following section, I will discuss an interesting possibility to use the machinery of generalized quantification outside the realm of structural semantics. I will appeal to ideas that I have presented in Verkuyl (2000) and more recently in Verkuyl (2013) in the hope to say a little bit more about associations, among which the two connected with Frans Zwarts. It is arguable that it is not Frans himself with whom I associate the two events but rather other situations in which Frans occurs. In the final section, associations will indeed be argued to be relations between situations. By so doing one can see that the phrase opening the present section—'When I think of Frans . . .'—is short for 'When I think of the situations in which I had something to do with Frans . . .'. Such an interpretation of that phrase opens the possibility of making two situations quite prominent within the set of all situations in which I had to do with Frans. Two situations shedding light on an interesting feature of his personality: his urge to evade negative reactions of his friends to making his own agenda too busy, which is too busy in order to make up arrears so as to evade negative reactions of his friends.

2. Generalized quantification and the lexicon

I assume some acquaintance with the marvelous work on Generalized Quantification done by Frans in Zwarts (1981) and Zwarts (1986). But I need to give some elementary stuff in order to show how lexical semantics can be involved in the machinery that was intended to do justice to the semantics of sentences.

So let us start with the basic model in Figure 1.

---

[2] For this distinction, see Church and Hanks (1990), Rapp (2002).

[3] See Washtell and Markert (2009). The relevant notion for distinguishing the distances between associated words is that of a *window* (a restricted room for measuring).

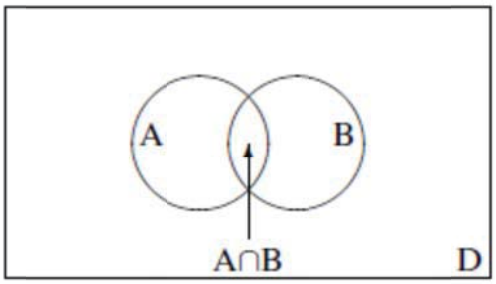

Figure 1: Intersection model of generalized quantification

In particular, determiners were discovered as the most important elements in a sentence because in sentences like (1a) analyzed syntactically in (1b):

(1) a. Some linguists in the Netherlands were rector magnificus.
b. DET (L,R)

they provide the proper quantificational information about the intersection L∩R, given that L (= A) pertains to the denotation L of the noun linguist and R (= B) denotes the VP be rector magnificus in the domain of interpretation D.

Skipping the account of tense, this yields the information that the intersection L ∩ R counts at least a member of the class that has the property of being a linguist and a rector magnificus. One may replace some in (1a) by all, which makes (1a) false, or by one, which makes it necessary to go into the history of the Dutch universities, or by few which makes (1a) true, etc. So, generalized quantification is about all, one, some, the, few, etc.

Let us also have a look at sentences like (2a).

(2) a. Some penguins show a high degree of pair fidelity.
b. SOME(P,F)

Here, there are two possibilities. The first one is that speaker and hearer know, say by being present in a situation in which they see penguins and talk about them, that (2a) is about the penguins in this domain of interpretation and so Figure 1 does what it is supposed to do: the sentence is true if there are F-penguins in the intersection P ∩ F. However, the sentence can also be used in a conversation in order to differentiate between sorts of penguins, say Humboldt penguins (+F) and emperor penguins (-F). In that use of (2a) the speaker and hearer do not have a particular domain D in mind: (2a)

takes into account all possible domains of interpretation in which the noun penguin applies to members of the penguin species.

In the two cases just illustrated by (1) and (2) the two sentences are structurally analyzed in terms of taking the determiner of the NPs some linguists and some penguins as providing the intersection A ∩ B, where the generalization over a set of domains D in (2) is to be taken as being due to other factors, such as the possibility of NPs to pertain to both types and tokens.

In his analysis of the semantics of crossword puzzles, Verkuyl (2014) discusses four ways in which a clue *C* may relate to and answer *A* on the basis of a set-theoretical relation between their denotata C and A, respectively.

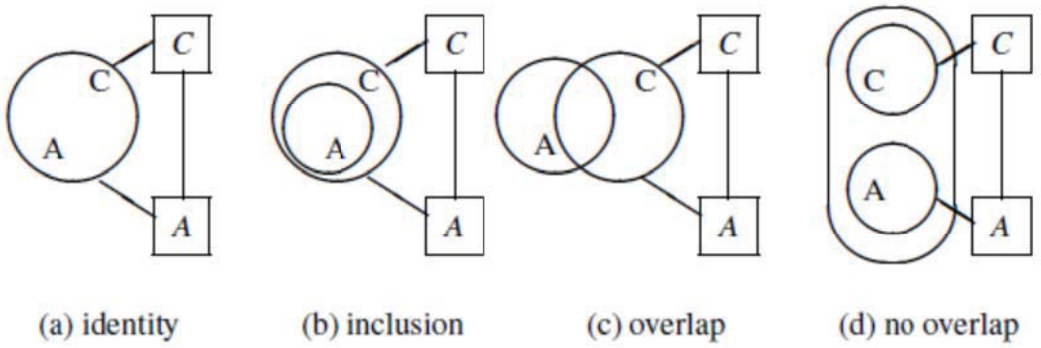

Figure 2: Four ways of relating lexical meaning

Figure 2 speaks for itself: in (a) the clue *C* and its answer *A* are or contain synonyms (e.g. *C = shut, A = close*), in (b) *A* is a hyponym of *C* because the set A is a proper subset of C (e.g. *C = seabird, A = penguin*), in (c) the clues A and C do not share a meaning element because there is just some overlap (e.g. *C = animal with pair fidelity*, *A = penguin*), and finally in (d) one might think of cases where *A = off* and *C = on*, or where *A = man* and *C = woman*.[4]

The configurations of the sets A and C in Figure 2 are familiar in the Theory of Generalized Quantification, but there are two obvious differences: (a) there is no determiner around in order to connect the sets A and C so as to get information about their intersection; (b) the four relations between A and C do not concern one domain of

[4] It is important to maintain the difference between italics and roman: clues and answers belong to a language L and so they are italicized. Their denotation occurs in roman capitals representing sets in domains of interpretation. It should also be clear that in (d) some sort of negation element is to be assumed in either answer *A* or clue *C*.

interpretation but all domains of interpretation available.[5] In all possible worlds accessible to the puzzle solver penguins are seabirds (P ⊂ S), irrespective of the individual members of P or S in a specific domain of interpretation and in some worlds all penguins may have the property of pair fidelity but this is not the case in all words. In general, the full meaning of a word A can be seen as pertaining to a set of sets C1 , C2 , . . . Cn containing A, much in the way proper names are often seen in terms of all sets of which its bearer is a member.

Breaking away from the model-theoretic position in which the interpretation of a certain set leads to some value in an objective domain of discourse "out there", one may use the machinery of formal semantics for modelling our cognitive capacity of dealing with meaning. That is, one may see the configurations in Figure 2 as flexible in the sense that individual speakers are always prepared to adapt the meaning of the words belonging to their mental lexicon on the ground of their experience. Ignoring the puzzle context of Figure 2, Verkuyl (2009; 2013; 2000) underscore the willingness of people to empty the A – C part of configuration (c) —that is, that part of A containing no members of C—so as to obtain configuration (b). This is called stereotyping. Against better judgment, the speaker ignores possible worlds in which the configurations (c) and (d) hold in favor of assuming (b) for all worlds. Think of word pairs like:

- man (A) — strong (C)
- jesuit (A) — unreliable (C)
- blond (A) — simple minded (C)
- model (A) — anorexic (C)

Etcetera, etcetera. It should be clear that in all these cases configuration (c) applies for the speakers of a language community and that stereotyping the lexical relations exemplified here leads to accepting (b) as the proper way to go for individual speakers, individually or as member of a larger group of that community. In this way, the covert "determiner" is promoted to the role of a universal quantifier. This seems to me the essence of stereotyping from the semantic point of view.

[5] Technically we have to assume a function *h* such that for all *w* in the set of possible worlds *W* : *h(w)* is the denotation of *A* in *w*. Assuming this for *C* too, the four configurations in Figure 2 can be seen as defining lexical relations between *A* and *C* by assuming that in all worlds *w*, A = C in (a), A ∩ C = A in (b), A ∩ C = ∅ in (c) and A ∩ C = ∅ in (d).

3. Associations and generalized quantification

As suggested earlier the notion of association is quite complex in the sense that the main divide between paradigmatic and syntagmatic associations discussed above gives room to a wide variety of sorts. As observed earlier they are generally not taken as meaning relations and in many cases it is even impossible to speak of an overlap: the intersection between the set of Fridays and the set of fishes is empty and there is no way to connect Frans Zwarts with faxes or beeps or post strikes in terms of a possible overlap. This could be an indication for bringing associations outside the domain of lexical semantics.

Yet there are several reasons for steering a different course. One of the problems for finding a unifying principle that covers all the different sorts of association is that in the literature the associative relation A(X,Y) is mainly defined in terms of X and Y as standing for sets of objects (I ignore here mass nouns). This problem may be solved by assuming that there is a binding factor which provides the possibility for the Theory of Generalized Quantification to enter the field. People generally assume a relation between Proust's madeleine (an object) and a memory of a situation in order to be able to use the term association for this connection. But one could also say that it is the situation in which Marcel smells his madeleine and thinks back to situations in his youth in which the madeleine also played a role.

On this account it becomes possible to make use of the four logical possibilities of Figure 2. And indeed, it is not too difficult to bring associations within the bounds of generalized quantification. Skipping the puzzle part of Figure 2, associations can now be seen, in Figure 3, as possible relations between situations. Possibility (a) could count as covering the associations of Pavlov's dog, whereas in (b) the set C contains all the situations in which I had to do with Frans Zwarts and A contains the two situations described at the beginning of the present chapter.

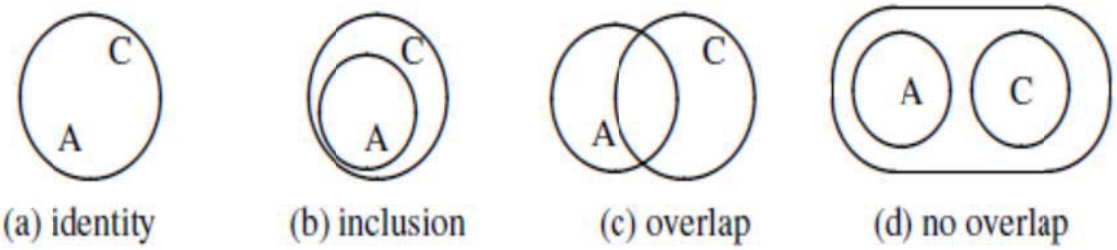

Figure 3: Four logical possibilities of relating sets of situations

The Friday-fish-example seems to fall under (c) with C for the situations in which people ate fish and A for situations in which it was dinner time on Friday. For centuries, in European countries the name Friday was associated with the word fish: most of the people ate fish on Friday, especially in catholic families. It is reasonable to assume that in spite of this habit people did not always eat fish on Friday and that fish was also eaten on other days of the week. Configuration (d) gives room to dreams.

It is not my purpose to present here a generalized quantified account of associations. This is just an initial impetus, so I restrict myself here to drawing Frans' attention to this perhaps unexpected extension of a field in which he was so adroit and virtuoso and which brought him his professorship in Groningen. Who does not remember his work on anti-additivity, antimorphisms, proper quasi-filters, filters and ideals? So, given the enormous amount of time available after his retirement, it would be a nice outlook for the two of us to work on a joint paper about the modeling of strength in associative relations. The only problem seems to be that I no longer have a fax ...

4. References

Church, K. W. and P. Hanks (1990). Word association norms, mutual information and lexicography. *Computational Linguistics* 16, 22–29.

Greenwald, A. G., D. E. McGhee, and J. L. K. Schwartz (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology* 6, 1464–1480.

Greenwald, A. G., L. Rudman, B. Nosek, M. R. Banaji, S. D. Farnham, and D. S. Mellott (2002). A unified theory of implicit attitudes. *Psychological Review* 109, 3–25.

Greenwald, A. G. and N. Sriram (2010). No measure is perfect, but some measures can be quite useful. Response to two comments on the brief implicit association test. *Experimental Psychology* 57, 238–242.

Rapp, R. (2002). The computation of word associations: Comparing syntagmatic and paradigmatic approaches. *Proceedings of the 19th International Conference of Computational Linguistics* 1, 1–7.

Verkuyl, H. J. (2000). Stereotyping, prototyping, and figurative use: Towards a proper semantic analysis. In T. F. Shannon and J. P. Snapper (Eds.), *The Berkeley Conference on Dutch Linguistics 1997. The Dutch Language at the Millennium*, 21–43. Lanham, New York, Oxford: University Press of America.

Verkuyl, H. J. (2009). Metataligheid als lexicografisch probleem. In E. Beijk, L. Colman, M. Göbel, F. Heyvaert, T. Schoonheim, R. Tempelaars, & V. Waszink (Eds.), *Fons Verborum.* Feestbundel Fons Moerdijk, 97–106. Leiden: INL.

Verkuyl, H. J. (2013). Propositional attitudes, stereotyping and the mental lexicon. In M. Aloni, M. Franke, and F. Roelofsen (Eds.), *The dynamic, inquisitive, and visionary life of ϕ, ?ϕ, and ◊ϕ. A Festschrift for Jeroen Groenendijk, Martin Stokhof and Frank Veltman*, 265–273. Amsterdam: ILLC.

Verkuyl, H. J. (2014). Word puzzles. In R. Taylor (Ed.), *The Oxford Handbook of the Word*, pp. 25. Oxford: Oxford University Press, to appear.

Washtell, J. and K. Markert (2009). A comparison of windowless and window- based computational association measures as predictors of syntagmatic human associations. *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing* 10, 628–637.

Zwarts, F. (1981). Negatief polaire uitdrukkingen I. *Glot* 4-1, 35–132.

Zwarts, F. (1986). *Categoriale grammatica en algebraïsche semantiek*. Ph.D. thesis, Rijksuniversiteit Groningen, Groningen.