What does Typicality stand for?

A formal analysis of Typicality as an effect of the gradual learning about members in the denotation

LOFT6, 2004 July 17th

Galit W. Sassoon,
The linguistics department,
Tel Aviv University

gala@post.tau.ac.il

Abstract

Typicality is extensively investigated in a variety of disciplines (Philosophy, Psychology, Linguistics, AI, etc.) However, despite the wide range of evidence for the existence of typicality effects, it is still not clear what the correct account for them is and how they relate to predicate meaning. I will present here Partee & Kamp's 1995 influential model theoretic analysis of typicality, which is (like most accounts of typicality) formulated in terms of predicate prototypes and degree functions. I will point out several problems with this kind of formulation (such as wrong predictions, compositionality problems, the complicated taxonomy of predicate types assumed by the theory etc.) Finally, I will attempt to show that they are all solved, if a typicality ordering of a predicate P in a context c ($\leq$P) is taken to reflect gradual addition of entities into the set of contextually relevant Ps, $[P]^{+}_{c}$ (and non-Ps, $[P]^{-}_{c}$), as information accumulates.

I. The basic facts

1. *Typicality scales* ($[\leq P]_c$): For any predicate P in a context c, speakers have intuitions about ordering of entities as to how *typical they are of P*. Evidence was first found by Rosch 1973 (following Witgenstein 1953) and was replicated time and again ever since. For ex. a robin, d_1, is usually considered *more typical of a bird* than an ostrich, d_2 ($\langle d_2,d_1\rangle \in [<$ *(typical-of a) bird*$]_c$).

2. *Typicality properties* ($OS^+_{(P,c)}$): Speakers have intuitions about a set ('cluster') of properties, which help to order entities by *typicality relative to P*. E.g. a robin scores better than an ostrich relative to: {*small*, *flies*, *sings, builds nests, perches in trees,* …} i.e. the cluster of ordering criteria of *bird* – $OS^+_{(bird,c)}$.[1]

3. It is easier to determine the P-hood of typical entities, even unconciously: e.g. the verification time of *a robin is a bird* is faster than that of *an ostrich is a bird* (Rosch 1973, Armstrong et al 1983).

4. Typical instances are acquired earlier (first) by children and adults (Rosch 1973, Smith 1988).

5. Typical instances are remembered best (Heit 1997).

6. They are retrived faster (Rosch 1973, Batting & Montague 1969).

7. In turn, typical (and early acquired) instances play a role in future encoding in memory, of entities and properties (Heit 1997, Rips 1975, Osherson 1990).[2]

8. Typicality effects occur with every predicate: People even order entities like *even* or *odd number*s by typicality and verification time for the more typical entities is faster (Armstrong et al 1983).

For ex. 5973954 is often considered an atypical *even number* (due to the *odd digits*, Wanner 1979).[3]

9. The typicality effects are highly context dependent: Within a context such as: *the bird walked across the barnyard*, a *chicken* is regarded as a typical bird, and verification time is faster for *chicken*, than for *robin* (Roth and Shoben 1983).[4]

II. Supermodels, Partee & Kamp 1995 (P & K):

[1] A supermodel M^* for a set of predicates A and a set of entities D is a tuple <M,T,m> s.t.:

1. **M is a partial model**: predicates are associated with partial denotations in M, $<[P]^+_M,[P]^-_M>$.

   For ex. if $[chair]^+_M = \{d_1\}$, $[chair]^-_M = \{d_3\}$, d_2 is in the gap: we don't yet know if its a chair or not.

2. **T is a set of total models**: completions of M (van Fraassen 1969): In each $t\in T$: denotations are total.

   $\forall t\in T, \forall P\in A$: 2.1. $[P]^+_t \cup [P]^-_t = D$; 2.2. $[P]^+_M \subseteq [P]^+_t$; 2.3. $[P]^-_M \subseteq [P]^-_t$.

   E.g. in each $t\in T$, d_2 is added to $[chair]^+_t$ or $[chair]^-_t$.

3. **m is a measure function** from sets of total models to real numbers between 0 and 1, s.t.:

   3.1. $m(T) = 1$; 3.2. $m(\{\}) = 0$; 3.3. $\forall T_1,T_2$, s.t. $T_1\subset T_2$: $m(T_1) < m(T_2)$.

[2] A predicates P is associated with a tuple $<p, c_e, c_P>$ s.t.:

1. **p is the prototype** – the best possible P.
2. $\mathbf{c_{e(d,P)}}$, is d's membership-degree relative to P: the degree to which d is P: $c_{e(d,P)} = m(\{t\in T: d\in[P]^+_t\}$.

   E.g. $1 = c_{e(d1,chair)} > c_{e(d2,chair)} > c_{e(d3,chair)} = 0$.
3. $\mathbf{c_{P(d,P)}}$ is d's typicality-degree relative to P: d's distance from P's prototype. Generally: $c_P \cong c_e$.

   E.g. in *chair*: the more typical entities fall under $[chair]^+$ in more of the total models.

[3] P & K distinguish between different predicate types in the following ways:

1. **+/– Vague:** The denotations of non-vague predicates, like *bird*, are total already in M.
2. **+/– Prototype:** *tall, odd number*, etc. have no prototype (there is no maximal *tallness* / *oddness*).
3. **+/– Typicality-is-coupled-with-membersip:** $\mathbf{c_P \cong c_e}$ (or: "the-prototype-affects-the-denotation")

   In *bird*, *grandmother* etc. typicality and membership are separated, because intuitively:

(1) A *pelican* d is a *bird* even in M (i.e. $c_{e(d,bird)} = 1$), but it is an atypical *bird* ($c_{p(d,bird)} < 1$). Thus, $\mathbf{c_e \neq c_p}$.

(2) $\forall d$, $\{t\in T: d\in[pelican]^+_t\}$ is a subset of $\{t\in T: d\in[bird]^+_t\}$, so $c_{e(d,pelican)}$ is smaller than $c_{e(d,bird)}$. But: d can be *more typical* of *a pelican* than of a *bird*, so $c_{P(d,pelican)}$ is bigger than $c_{P(d,bird)}$. Again, $\mathbf{c_e \neq c_p}$.[5]

P & K's influential analyisis considerably improved upon previous accounts based on Fuzzy logics (Osherson and Smith 1981). However, it still maintains some standard, but problematic, assumptions.

III. Problems:

A1. Since for them $c_e \neq c_P$, P & K do not say how c_p values (or typicality judgments) are determined.

A2. Degrees are not intuitive primitives (P & K): people tend to have natural intuitions about entities being more (or less) typical than others, but not about entities' exact numerical degree (say – 0.695).

A3. The separation between c_e and c_p (in predicates like *bird*) forces us into a non-elegant theory, that stipulates as primitives two non connected sets of values for c_e and c_p.

B1. c_e and c_P are total functions, **but speakers' knowledge about the typicality scale is not total:**

For ex. if one bird sings and the other flies, which one is *more typical*?

(Nor do speakers know every typicality property. For ex. is *in the home* typical of *chairs*?)

B2. The denotations of non-vague predicates (e.g. *bird*) are represented as total (Fregean entities independent of speakers' experience or belief). But, as noted above, typicality is connected to the set of entities, which a speaker knows and considers as relevant in the context.

C. Sometimes P & K cannot associate a prototype with predicates displaying typicality effects:

(1) Intuitivly, some people are *typical tall persons*. Since there is no prototype this is not accounted for.

(2) Similarly for *even / odd number*.

D. Problems of compositionality: To give one example, what would the prototype for negations, like *not a bird*, be, given the variety of possible *non-birds*: A dog? A pen? The number 52?

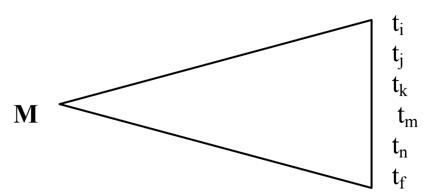
E. **We are forced into a complicated taxonomy of predicate types**, which is especially problematic when compositionality is addressed (of what type is a conjunction of different predicate types?).[6]

III. My proposal: Dimension models

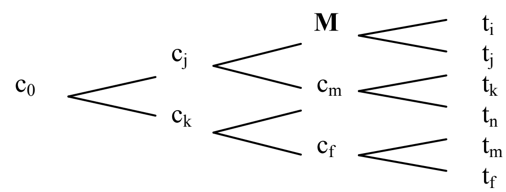
1. Replace <p, c_P, c_e> by the psychologically real set ('cluster') of typicality properties $OS^+_{(P,c)}$.
Due to time limits, I will not discuss OS today (but see the gray definitions below and notes 9-11).

2. Replace the partial model M with a whole set C of partial models ('contexts'), partially ordered by a relation $\leq$ representing monotonic information growth, i.e. context c_2 extends context c_1 ($c_1 \leq c_2$) iff the denotations in c_2 are supersets of the denotations in c_1. Denotations are empty in the minimal context c_0, and total in total contexts $t \in T$ (Veltman 1984; Landman 1999; etc.) Similarly, for clusters.

P & K:



My proposal:



A dimension model M* for a set of predicates A and domain D is a tuple $<C, \leq, c_0, T>$ s.t.:

[1] C is a set of partial models c ('contexts'): where any predicate P is associated with:

1. Partial positive and negative denotations: $<[P]^+_c, [P]^-_c>$

2. Partial sets ('clusters') of typical and non-typical properties (predicates): $<OS^+_{(P,c)}, OS^-_{(P,c)}>$

Ex.: if $OS^+_{(chair,c)}$ = {*4 legs, a back*}, $OS^-_{(chair,c)}$ = {pink}, *wooden* is in the gap: we don't yet know if it is typical of *chair* or not. The model applies to any n-place predicate and its OS dimensions, taken to denote sets of n-tuples of entities.

[2] $\leq$ is a partial order on C: $\forall P \in A$:

1. Monotonicity: $\forall c_1, c_2 \in C$, s.t. $c_1 \leq c_2$: $[P]^+_{c1} \subseteq [P]^+_{c2}$; $OS^+_{(P,c1)} \subseteq OS^+_{(P,c2)}$
$[P]^-_{c1} \subseteq [P]^-_{c2}$; $OS^-_{(P,c1)} \subseteq OS^-_{(P,c2)}$

2. c_0 is the minimal element in C under $\leq$: $[P]^+_{c0} = [P]^-_{c0} = \varnothing$; $OS^+_{(P,c0)} = OS^-_{(P,c0)} = \varnothing$
(Denotations and clusters are empty in the minimal context c_0).

3. T is the set of maximal elements under $\leq$: $[P]^+_t \cup [P]^-_t = D$; $OS^+_{(P,t)} \cup OS^-_{(P,t)} = A$
(Denotations and clusters are maximal (no gaps) in contexts t in T).

4. Totality: $\forall c \in C, \exists t \in T: c \leq t$.
(Every context has some maximal extension).

5. "Super-denotations": $[P]_c = \cap\{[P]^+_t | t \in T, c \leq t\}$; $[\neg P]_c = \cap\{[P]^-_t | t \in T, c \leq t\}$

3. Given the empirical data in section I, I propose that P's typicality scale in t denotes the order in which we learn that entities fall under P (or ¬P) in the partial contexts under t:

[3] $\forall t \in T$: $(<d_1, d_2> \in [\leq P]^+_t)$ **iff**: $\forall c \leq t$: $(d_1 \in [P]_c \rightarrow d_2 \in [P]_c)$ & $(d_2 \in [\neg P]_c \rightarrow d_1 \in [\neg P]_c)$.

(In any total t, d_1 is *equally or less (typical of) P* than d_2 iff in any c under t, if d_1 is P – d_2 is P and if d_2 is non-P – d_1 is non-P).

IV. Problems A-E can be solved now:

A1. We have pointed out how typicality judgments are determined. We correctly predict that the earlier we learn that an entity is P, the more typical P we consider it to be.

A2. Degrees are not directly given, but can be derived: $c_{e(d1,P,t)} \leq c_{e(d2,P,t)}$ iff $<d_1,d_2> \in [\leq P]^+_t$.

A3. Typicality is always coupled with membership ($c_p = c_e$), because, e.g.:

(1) Entities added in the denotation $[bird]^+$ in different contexts under t, recieve different degrees, so even if a *pelican* d is known to be a *bird* in c, it does not necessarily receive degree 1 (for $c_{e(d,\, bird)}$).

(2) An entity d can be added in $[bird]^+$ in extended contexts relative to other bird types (i.e. be an atypical *bird*) but still be the first *pelican* ever seen (i.e. the most typical *pelican* in the context).

(We captures the so called *conjunction effect / fallacy*).[5, 7]

B. The typicality scales are possibly partial, like any predicate denotation: Scales are given in each total context separately. A pair $<d_1,d_2>$ is in the gap of a scale $\leq$P in c, if c may still extend to totality with $<d_1,d_2>$ either in the positive ($[\leq P]^+_{t1}$, $c \leq t_1$) or in the negative scale ($[\leq P]^-_{t2} = [>P]^+_{t2}$, $c \leq t_2$).[8]

C. All the typicality effects are accounted for:

(1) Context-INdependent upper bounds (maximal tallness) are not needed. In each context c some entities are the best Ps (the *tallest*).

(2) Since $[even\ number]^+_c$ represents the set of *contextually relevant even numbers*, naturally, a gap exists and classification is gradual. E.g. the contextual relevance of a perfect *even-number* like 579,936 may be questioned (and quite often, relative to, say – 2 or 200,000).

D. As for negations, the scales of P and ¬P are predicted to be inverse: $\forall t$: $[\leq P]^+_t = [\geq \neg P]^+_t$, which seems intuitive (work in progress; see also Smith et all 1988, p. 511).

E. The taxonomy of predicates is unified. **A context dependent scale is indicated by the operator *typical of*:** For ex. in *taller* the ordering criterion, and hence the scale, is fixed semantically, but in *more typical of a tall person / chair / bird* etc., the ordering criteria, and scales, are context dependent.[9, 10, 11]

☺

Notes

1. OS^+ properties are fixed by language (e.g. *animate* for *bird*), world knowledge (*feathered, eats insects*), or context (*sweat / sour, red* for *fruit*, 2 / *4 wheels* for *tractor*). Very few or no $OS^+_{(P,c)}$ properties are necessary for membership in $[P]_c$.

2. Typical instances affect future encoding in memory of entities and properties. For ex.:
(1) Only the properties of typical instances are encoded as being related to the predicate (typical of it).
(2) Rips 1975, Osherson 1990 have shown that the *properties* of the typical instances are grasped as more likely to hold of atypical members than v.v. E.g. the inference: *if robins are susceptible to a certain disease then ostriches are susceptible to it*, is uaually stronger than the same inference from ostriches to robins.
(3) Speakers tend to misinterpret observed denotation members which are not satisfying the typical properties, as non-members or as members satisfying these properties (Heit 1997).

3. Typicality effects occur with <u>every</u> predicate**,** whether denoting actions, events, emotions, social and psycho-diagnostic concepts, natural kinds, programming concepts, or even mathematical concepts (Wanner 1979, Armstrong et al 1983).

4. Typicality affects a variety of linguistic phenomena / puzzles. For example:
(1) Lasersohn 1998 analyses <u>pragmatic looseness</u> (the possibility to interpret any statement as *true enough* when it is strictly false, if it comes *close enough to the truth for all "contextual practical purposes*"). Lasersohn associates with each predicate a contextual “halo” that contains sets that differ from the actual denotation only in <u>*pragmatically ignorable features*</u>, or in practice, typicality properties.
(2) It is standardly assumed that <u>domains of quantifiers and modals are contextually restricted by sets of properties</u> (see Von Fintel 1994 re. conditionals; Kadmon and Landman 1993 re. indefinite generics, and re. the proposal that *any* uses the restricting properties also in a unique way, i.e. in order to <u>widen</u> rather than to restrict; In my MA thesis, I have proposed that sometimes quantifyers like *any* make use of the typicality properties, so as to order, or to reduce the ordering, of the entities in the domain).
(3) Lewis 1970,1979, among others, analyses <u>gradeable predicates</u> (like *tall*, *cold*) and comparative adjectives (*cooler*, *taller*, or in general the structure *more P than*) using contextual features. In order to fit in this construction, non scalar predicates need to be modified like in: *more typical (of a) P*. Kamp 1975 suggests that nouns like *chair / bird* are usually non-gradable because they are assocciated with a set of ordering criteria (i.e. the typicality properties), not <u>one</u> unique criteria.

5. The conjunction effect / fallacy:
The intuition in (2) (that a pelican d can be more typical of a *pelican* than of a *bird*) is an istance of what is usually called *the conjunction effect* (Smith et al 1988): Take for ex. *stiped apple* vs. *apple*. Intuitively, a striped-apple d is regarded as more typical of *a striped apple* than of an *apple*. This has to be represented by c_p (the typicality degreee function).
A parallel effect, represented in P & K by c_e (the membership degreee function), is usually called *the conjunction fallacy* (Tversky & Kahneman 1983): e.g. a *striped-apple* d is sometimes judged as more likely *a striped apple* than an *apple* (despite the fact that the probability of being a *striped apple* <u>cannot</u> be greater than the probability of being an *apple*).

P&K's account for the so called conjunction fallacy:
P & K observe that modifiers (*striped*) receive a distinct interpretation in the local context created by the noun (*apple, zebra etc.*) Thus, they propose to replace c_e by a new function $c_{e(d,\, striped\, /apple)}$, which may assign d a higher value than $c_{e(d,apple)}$ or $c_{e(d,striped)}$:

$$c_e(d, striped\, /apple) \;=\; \frac{c_e(d, striped) \;-\; \min\{c_e(x, striped) \mid x \in [apple]^+_M\}}{\max\{c_e(x, striped) \mid x \in [apple]^+_M\} \;-\; \min\{c_e(x, striped) \mid x \in [apple]^+_M\}}$$

<u>The problem</u>: This doesn't work as an account for the so-called conjunction effect / fallacy in lexical nouns without a modifier, like *pelican* vs. *bird*. Given that complex phrases like *male-nurse* in English are simple lexical items in other languages (e.g. *male nurse* in Hebrew is: *ax*), we would prefer a unified account for both types of noun phrases.

6. To give one example, the predicate *bird* is sharp and has a prototype that doesn't affect the denotation. The predicate *tall* is vague and has no prototype. How do we construct meaning for *(an ostrich is a) tall bird*?

7. Similarly, d can be added in $[apple]^+_c$ in extended contexts relative to other apples (i.e. be an atypical *apple*), but still be the first known *striped apple* (the best *striped apple* in the context).

8. $\forall c \in C$: **(**$\langle d_1,d_2\rangle \in [\leq P]^+_c$**)** <u>**iff**</u>**:** $\forall t \in T, c \leq t$: **(**$\langle d_1,d_2\rangle \in [\leq P]^+_t$**).**
(In a partial c d_1 is *equally or less (typical of) P* than d_2 <u>iff</u> in every total t extending c d_1 is *equally or less (typical of) P* than d_2).

9. The typicality properties help to order entities on the P scale:

[4] $\forall t \in T$: $\boxed{[\leq P]^{+}_{t} = [\leq(\wedge OS^{+}_{(P,t)})]^{+}_{t}}$ I.e. d_1 is *equally or less typical of P* than d_2 iff d_1 is *equally or less typical* of (the conjunction of) all P's typicality properties ($\wedge OS^{+}_{(P,t)}$).

Given the definition in [3] the following intuitive results follow from [4]:

(1) If d_2 is *at least as typical as* d_1 relative to every $OS^{+}_{(P,c)}$ property, then d_2 is *at least as typical* P *as* d_1.

(because: $\cap\{[\leq Q]^{+}_{t} \mid Q \in OS^{+}_{(P,t)}\} \subseteq_{(by\ [3])} [\leq(\wedge OS^{+}_{(P,t)})]^{+}_{t} =_{(by\ [4])} [\leq P]^{+}_{t}$).

For ex. if *fly* and *sing* order birds, and d_1 is equally or less typical of *fly* than d_2, but more typical of *sing*, we cannot automatically determine which bird is more typical in c. But if d_2 is equally or more typical of *fly, sing* etc. (no inverse relations relative to different OS^{+} properties) we must see it as more typical of a *bird*.

(2) If v.v, i.e. they are *equally typical* relative to every $OS^{+}_{(P,c)}$ property (*fly, sing, perch* etc), they are *equally typical* Ps (*birds*)

(because: $\cap\{[=Q]^{+}_{t} \mid Q \in OS^{+}_{(P,t)}\} \subseteq_{(by\ [4])} [\leq P]^{+}_{t} \cap [\geq P]^{+}_{t} = [=P]^{+}_{t}$.).

(3) The ordering of entities in inverse ordering relations relative to different typicality properties is more indexical in nature, and not always known. However, intuitively, we may regard d_2 as *more typical* of a *bird* than d_1, even if say – d_2 is more typical in *singing* but less typical in *flying*, providing that it is d_2 that is, overall, more typical relative to *flying and singing* (because say – *singing* is more crucial with respect to birds in the context). In any event, d_2 must be better relaive to at least one $OS^{+}_{(P,t)}$ property.

(because: $[\leq P]^{+}_{t} =_{(by\ [4])} [\leq(\wedge OS^{+}_{(P,t)})]^{+}_{t} \subseteq_{(by\ [3])} \cup\{[\leq Q]^{+}_{t} \mid Q \in OS^{+}_{(P,t)}\}$).

10. General Conclusions: Towards a better representation of predicate meaning

The analysis presented here is but one case study (among several others, Sassoon 2001,2002) demonstrating the usefulness of a vagueness model in which the sets of typicality properties of a predicate are explicitly represented. This model allows for a fuller account of people's knowledge of predicate meaning, the way they acquire it, and a variety of psychological findings.

In addition it allows for a systematic account for a variety of semantic phenomena. In model theoretic semantics, clusters of properties /dimensions /features /propositions etc. are used in the analysis of a variety of phenomena (see for instance, Lewis 1970,1979; Kamp 1975; Bartsch 1984-1998; Landman 1989; Kadmon and Landman 1993; Lasersohn 1998; etc.) However, there is no systematic account for the role of clusters. Stipulations are made in each case separately, regarding their presence or effects. In fact, the denotational theory came into being partially as a reaction against cluster theories of meaning:

[1] These theories did not account for the relation of words to non-linguistic entities (Lewis 1970);

[2] They did not account for the composition of complex phrases' meaning from their parts' meanings (Osherson et al 1981);

[3] They presented a static picture of total information about predicate meaning that did not capture cases of ignorance, gradual learning, error etc. (Laurence and Margolis 2003);

[4] They had some crucially wrong predictions (Kripke 1972, Putnam 1975).

My model overcomes these deficiencies:

[1] It implements OS^{+} within a denotational theory, and the connections between OS^{+} and the denotation are properly defined[9];

[2] It seems to be more fruitful in dealing with compositionality (a small part of my compositionality account is in section IV.D (work in progress); see related work also in Hampton 1997);

[3] As a vagueness model, it allows for ignorance, gradual learning etc.;

[4] It overcomes the classical Kripke-Putnam objections. For ex. in certain, but not all contexts, the properties *mammal* or *living in the North Sea* are taken to be presupposed by *whale*. So *whale* does not entail its cluster properties, and it doesn't follow that *whales are mammals* is a tautology, or *whales are fish* a contradiction.

11. The analysis extends even to non-intersective modifiers: For ex. the predicate *stone-lion* associates with the set of stones which, in the context, are typical enough relative to $OS^{+}_{(\mathbf{lion},t)}$ ($OS^{+}_{(lion,t)}$ orders lions as well as non-lions, including stones). The standard of precision varies from context to context (Lewis 1970, 1979), but at least we can confidently assume, like in the case of *tall*, that *stone-lion* associates with the most typical stones on the typicality scale of *lion*.

References

1. Armstrong, S.L.,Gleitman L.R. and Gleitman H., 1983, "What some concepts might not be", *Cognition* 13, 263-308.
2. Bartsch, Renate, 1986, "Context dependent interpretations of lexical items", J. Groenendijk, D. de Jongh, M. Stokhof (eds.) *Foundations of Pragmatics and Lexical semantics*, GRASS 7, Foris, Dordrecht.
3. Bartsch, Renate, 1984, "The structure of word meanings: Polysemy, Metaphor, Metonymy" in Landman Fred & Veltman Frank (Eds.), *Varieties of Formal semantics*, GRASS 3, Foris, Dordrecht.
4. Bartsch, Renate, 1998, *Dynamic Conceptual Semantics*, CSLI publications, Stanford, U.S.A.
5. Batting W.F. & Montague W.E., 1969, "Category Norms For Verbal Items In 56 Categories" *Journal Of Experimental Psychology Monograph* 80 (3) Pt. 2.
6. Hampton, James, A, 1997, "Conceptual Combination", K.Lamberts & D.Shjanks (Eds) *Knowledge, concepts and categories,* pp. 135-162, Psychology Press.
7. Heit, Evan, 1997, Knowledge And Concept Learning, K.Lamberts & D.Shjanks (Eds) *Knowledge, Concepts And Categories,* The MIT press, Cambridge, MA, 7-42.
8. Kadmon, Nirit, and Fred Landman, 1993, 'Any', *Linguistics and Philosophy* 16, 353-422.
9. Kamp, Hans, 1975, "Two theories about adjectives", E.L.Keenan (Ed.), *Formal Semantics of Natural Language,* 123-55. Cambridge University Press, Cambridge.
10. Kripke, S.A., 1972, "Naming and necessity", G.Harman and D.Davidson (Eds.), *Semantics of Natural Language,* Kluwer, Dordrecht.
11. Landman, Fred, 1989, "Groups II" in *Linguistics and Philosophy* 12, 723-744, Netherlands.
12. Landman, Fred, 1991, *Structures for Semantics*, Kluwer, Dordrecht.
13. Lasersohn, Peter, 1998, "Pragmatic Halos", ms, University of Illinois.
14. Laurence, Stephen and Eric Margolis, 2003, "Concepts", in: Ted Warfield & Stephen Stich (Eds.) *The Blackwell guide to the philosophy of mind,* Blackwell.
15. Lewis, David, 1970, "General Semantics", *Synthese* 22, 18-67. Reprinted in D.K. Lewis, (1983), *Philosophical Papers*, Vol. 1, 189-229, Oxford University Press, New York.
16. Lewis, David, 1979, "Scorekeeping in a Language Game", *Journal of Philosophical Logic* 8, 339-59. Reprinted in D.K. Lewis, 1983, *Philosophical Papers,* Vol. 1, 233-49, Oxford University Press, NY.
17. Osherson, D and E. Smith, 1981, "On the adequacy of prototype theory as a theory of concepts", *Cognition* 11, 237-62.
18. Osherson, D.N. And E.E. Smith, O.Wilkkie, A.Lopez, E.Shafir, 1990, "Category Based Induction" *Psychological Review* 97, 185-200.
19. Partee, Barbara & H. Kamp, 1995, "Prototype Theory and Compositionality",*Cognition* 57: 129-91.
20. Putnam, H., 1975, "The meaning of 'meaning'", Keith Gunerson (Ed.), *Minnesota Studies in the Philosophy of Science,* Vol. VII, University of Minnesota Press, Minneapolis.
21. Rips L.J., 1975, "Inductive Judgments about Natural Categories", *Jurnal of Verbal Learning and Verbal Behaviour* 14, 665-81.
22. Rosch Eleanor, 1973 "On The Internal Structure Of Perceptual And Semantic Categories", In T.E. Moore (Ed.), *Cognitive Development And The Acquisition Of Language*, New York, Academic Press.
23. Roth E.M. & E. Shoben, 1983, "The Effect of Context on the Structure of Categories", *Cognitive Psychology* 15, 346-78.
24. Sassoon, Galit, 2001, "Predicate interpretations as intensions restricted along dimensions", MA thesis, Tel Aviv University, Israel.
25. Sassoon, Galit, 2002, "Semantics with clusters of properties", in: Yehuda Falk (Ed.) *Proceedings of Israel Association for Theoretical Linguistics 18*, Bar Ilan University.
26. Smith, Edward E., 1988, "Concepts and thought", *The Psychology of Human Thought,* Cambridge University Press.
27. Smith, Edward E., D.N.Osherson, J.Rips, M.Keane, 1988, "Combining Prototypes, a Selective Modification Model", *Cognitive Science* 12, 485-527.
28. Tversky, Amos & D. Kahneman, 1983, "Extensional versus intuitive reasoning: The conjunction fallacy in probability judgments", *Psychological Review* 90, 293-315.
29. Van Fraassen, Bas C., 1969, "Presuppositions, supervaluations and free logic", K. Lambert (Ed.), *The Logical Way of Doing Things*, 67-91, New Haven, Yale University Press.
30. Veltman Frank, 1984, "Data semantics", Jeroen Groenendijk & T.M.V.Janssen & Martin Stokhof (Eds.) *Truth, interpretation and information* Proceedings of the third Amsterdam colloquium, 43-63.
31. Wanner E., 1979, "False Identification Of Prime Numbers", Paper Presented At The 1979 Meeting Of *The Society Of Philosophy And Psychology,* New York, NY.
32. Wittgenstein, Ludwig, 1953, *Philosophical Investigations,* translated by G.E.M Anscombe, third edition, 1968, Blackwell, Oxford.