Multilinear Semantics for Double-Jointed and Convex Coordinate Constructions

Arthur Merin
Department of Philosophy
University of Konstanz
2012

Abstract: An exemplary case study is offered for multilinear semantics [MLS] in linear spaces over orderable rings R, notably the real numbers $\mathbb{R}$.. In particular, English '*and*' and '*or*' denote operations that form R-linear combinations of vectors denoted by syntactic coordinables. Scalar coefficients are values of constant or context-dependent selection functions with characteristic constraints that distinguish distributing '*and*', '*or*', and non-distributing, convex combination '*and*'. Intercategorial semantic composition of intracategorially coordinables is either applicative or by tensor product. Either way, function/argument role distinctions have no empirical import in MLS. A comparison with Montague semantics and Keenan-Faltz boolean semantics shows many predictive advantages. A brief outline only is given of negation, which is represented at its most general as half-space formation induced by preference-valuations. MLS indeed replaces truth-values as generators of structure for meaning recursion by desiderative utility-values in R and, as a special case, evidential relevance values. This motivation is compared to that for recent vector space semantics (Preller & Sadrzadeh 2011), which are truth-theoretic

and utilize compact 2-category structure common to subspace projector lattices of real vector spaces, thence quantum logic, and two-sorted first-order logic. Phenomena involving empirical failures of distribution and absorption laws are noted whose explanation rests on R-linear and truth-conditional representations interacting so as to warrant a hypothesis of dual representation.

1 Multilinear Semantics

Multilinear Semantics [MLS] proposes denotata for natural languages in linear algebras over orderable number rings R (i.e. $\mathbb{Z}$, $\mathbb{Q}$, or $\mathbb{R}$). The most familiar models of such algebras are (i) the Euclidean spaces of idealized visual perception, images, and locomotion, and spaces of valuables such as (ii) commodity bundles and (iii) decision-theoretic acts construed as random variables. The latter two are familiar from economics (cf. Merin 1986, 1988, 1997, 2002). Semantics is for predicate-logic-sized fragments, and not confined to locatives (Zwarts and Winter 2000). MLS has certain advantages over logical or lattice-theoretic semantics in predicting spontaneous native speaker intuitions on acceptability and paraphrase. We take it to co-exist and to interact with truth-conditional or truth-related evidence-conditional construals of utterances or arguments, comparably to 'dual coding' hypotheses in cognitive psychology.

A simple way to approach MLS from ground familiar to meaning theory is to adopt, first, the tenets of the general parts (Secs. 1–3) of Montague's 'Universal Grammar' (Montague 1970) [UG]. In line with standard doctrine in computer science, they specify that a semantics is to be a homomorphism from a syntactic algebra to an algebra of denotata. The intent is to capture the pre-theoretical idea of compositionality.[1]

As of Sec. 4, Montague quite naturally postulates truth-values to constitute the space of (call it) 'ultimate values' which semantic valuations map to. MLS, by contrast, postulates use-values with ordered group structure familar from the integers $\mathbb{Z}$, rationals $\mathbb{Q}$ and reals $\mathbb{R}$—rather than truth-values. The laws of quantity (Hölder 1901) or their discrete restrictions to $\mathbb{Q}$ or $\mathbb{Z}$ apply to these spaces. The fundamental question, as it were, that informs the constitution of meaning is no longer "Is it so or not?', but rather "How much better or worse is it than the *status quo*?"

Under interpretation (ii/iii), MLS denotation algebras will relate to spaces of use-values much as boolean denotation algebras relate to the space of truth-values. The (classical) truth value space, **2**, is the simplest non-degenerate boolean algebra up to isomorphism. A use value space, representing preferences of the cardinal or yet more narrowly defined numerical scale type, is up to isomorphism the simplest nondegenerate linear algebra of the sentence denotation algebra type. For an $\mathbf{R}$-vector space, e.g. $\mathbb{R}^n$, this will be $\mathbf{R}$, here: $\mathbb{R}^1$, i.e. $\mathbb{R}$. Structure may be lifted 'pointwise' from this space of ultimate values.[2] Alternatively, sentence-denotata are conceived from the outset as vector elements of some linear space. They are then mapped to their ultimate value space by valuations, which in MLS will be linear functionals. Common to both linear, boolean, and conceivably other algebraic approaches to semantics is that valuations are morphisms from some algebra to the smallest nondegenerate algebra in the respective category.

Linear polynomial composition in MLS proceeds by *linear combination*. A linear combination of $n = 2$ elements $\mathbf{x}$, $\mathbf{y}$ of an R-linear space, $\mathcal{L}$, is an element $a\mathbf{x}+b\mathbf{y}$ of $\mathcal{L}$ where scalars a, b are elements of the underlying ring R that operate on the additive group of $\mathcal{L}$.[3] In any application, the a, b, may be values $a = \alpha_{\tilde{c}} = \alpha(\tilde{c})$, $b = \beta_{\tilde{c}} = \beta(\tilde{c})$ of functions α, β from some set $C = \{\tilde{c}, \tilde{c}', \ldots\}$ into R. Intuitively, the $\tilde{c}$ are conceived of as contexts-of-use for sentences. So utterances will be (sentence,context) pairs, with contexts, much as David Lewis suggested, conceived of as m-tuples of descriptive coordinates. In our applications, $1 \leq m \leq 3$.

Coordination in MLS and, by n-fold generalization to linear combinations $\sum_i a_i\mathbf{x}_i$, also finitary quantification make use of such indexed linear combination, and there will be infinitary generalizations. Hetero-categorial semantic composition is exemplified for string parse schemata NP·VP that stands for concatenation of a syntactic object of type noun phrase with one of type verb phrase. Such composition is either by (a) typed functional application $[\![\mathrm{VP}]\!]([\![\mathrm{NP}]\!])$ or $[\![\mathrm{NP}]\!]([\![\mathrm{VP}]\!])$ in which the function is a linear map on the linear space of which its argument is an element[4] or else by (b) tensor product $[\![\mathrm{NP}]\!] \otimes [\![\mathrm{VP}]\!]$ (see e.g. MacLane & Birkhoff 1967).

The tensor product construal in effect allows, though it does not enforce, a relaxation of the relatively tight relationship between semantic typing and inter-categorial syntactic typing that is found in Montague and similar applicative semantics (Merin 1988, 2002).[5] It is also less restrictive about possible interpretations of transcategorial semantic composition and decomposition than will be recursive applications of linear maps to vector objects. The latter may suggest interpretations that might be unexpectedly illuminating in some cases, but simply gratuitous in others.

A word token from {*and, or*} denotes a (binary) *indexed linear combinator*, i.e. a function $\lambda\tilde{c}\mathsf{xy}[\alpha_{\tilde{c}}\mathsf{x} + \beta_{\tilde{c}}\mathsf{y}] \in \mathcal{L} \times \mathcal{L} \rightarrow \mathcal{L}$ where $\mathcal{L}$ ranges over linear denotation spaces $[\![\mathrm{S}]\!]$, $[\![\mathrm{XP}]\!]$, ... of conjoinable syntactic categories S, XP, ... The set C has for elements index-tuples $\tilde{c} = \langle c^\circ, k, w\rangle$ of ⟨utterance-context, k-th occurrence-in-cotext, world⟩. Index k individuates tokens of a relevant indexical designator (in our binary case, a connective such as '*or*)' in the sentence or discourse; the w are intuitively as in familiar versions of intensional semantics; c° represents situational pragmatic features of the use-context not addressed by k and w. Thus, MLS hypothesizes for

- $[\![and]\!]_{[+D]}$ ('distributing'), $\alpha_{\tilde{c}} = \beta_{\tilde{c}} = 1$; i.e. simple vector addition, a non-idempotent operation;

- $[\![and]\!]_{[-D]}$ ('non-distributing, collective, convex'), $\beta_{\tilde{c}} = 1 - \alpha_{\tilde{c}} \in]0,1[$, conjoining sub-clausal constituents only;

- $[\![or]\!]$, $\beta_{\tilde{c}} = 1 - \alpha_{\tilde{c}} \in \{0,1\}$, such that α is a random variable ostensibly under the control of some agency other than the Speaker: (i) for imperatives, the addressee; (ii) for simple indicatives, Nature or a specifiable non-speaker party.

Negation and conditionals, the obvious candidates for completing the basic set of sentential operations, are not straightforward in MLS. They are thus not as obvious as the orthocomplement and appropriate residuation operations in subspace lattice-based approaches to implementing truth(-like) semantics in linear spaces. The conditional is not addressed at all in the present short article (see Merin 2002 for it and negation), and only a brief, discursive indication will be given in Section 5.1 of how negation is most generally introduced.

2 Empirical Motivations

Compositional, logical semantics (Montague 1973 [PTQ], Keenan & Faltz 1985 [KF]) has been overlooking elementary descriptive problems already for sentential languages. For instance, it mispredicts paraphrastic equivalence, intelligibility, and acceptability of would-be English instances of the (i) distributive, (ii) absorption and (iii) idempotent laws (Merin 1997, 2002). Appeal to a 'Maxim of Manner' (Grice 1967), specifically on grounds of unnecessary expression length, might be held to account for failures of (iii), but will do so unimpressively. Length does not explain why and how (a) '*Kim is at home now and Kim is at home now*' and (b) '*Kim is at home now or Kim is at home now*' differ in degree and kind of unacceptibility. Ex. (a) is plain weird, ex. (b) might be a smart aleck's facetious way of saying that Kim is at home now.

Even if 'Gricean implicature' rationales were found sufficient to explain the data on (iii), they could not explain deviations from the distributive and general lattice law predictions for (i) and (ii), where two out of four laws, one from each dual pair of laws, are seen to fail. MLS, by contrast, predicts the data under a dual-representation assumption. Specifically, it does so in concert with constraints on probabilistic evidential relevance valuations of atomic sentences (Merin 1997:§7, 2002, and see Section 5.2 below for (i)).

For subsentential coordination, PTQ and KF have problems which are rather more familiar from the literature. These problems have been addressed with partial success by a combination of type shifting regimes (PTQ, KF, ...) and of ontological hybridization of phrasal coordination structure. The latter is derived from boolean, sentential coordination by appeal to weaker semilattice structures defined on entities of lowest, 'individual' semantic type (Link 1983). A mature instance of the type-shifting technique is Winter (1998/2001).

Type shifting, with or without added ontological hybridization, is a powerful instrument for matching intuitions to putative logical form. It is thereby reminiscent of some of the 'syntactic transformations' of early Generative Grammar. Indeed, 'Conjunction Reduction' is a historic transformation that is motivated for proper name NPs by shifting their type from non-conjoinable type E (individual) to conjoinable type $[[E \rightarrow T] \rightarrow T]$, where T is the type of truth values, the Fregean extension of indicative sentences. The motivation for type shifting may in crucial cases simply be to obtain empirically correct distributing or non-distributing readings, which algebra-intrinsic constraints on boolean homomorphisms do not afford without it.[6]

The interest of MLS for the analysis of (non-)distribution under paraphrastic equivalence arises on *a priori* grounds. Spaces of homomorphisms of boolean algebras are not in general boolean algebras (see e.g. Keenan & Faltz 1985). By contrast, spaces of linear maps on linear algebras, i.e. spaces of linear homomorphisms, are again linear algebras.

It follows that there is no use for type shifting in MLS. Function/argument role allocation has no implications for distributivity of one coordinative semantic operation over another or for that of a phrase type instance over the coordinator of a concatenated phrase type instance. Specifically, it will have no such implications for word strings parsed as concatenations, NP·VP, of coordinate noun phrase and coordinate verb phrase, Accordingly, we compare predictive abilities and expenses of MLS with those of broadly boolean semantics augmented by type shifting and ontological hybridization facilities. We do this for a class of natural language constructions, here of English, which bring the issue to a fine, but readily scrutable point.

Both linearity and the indexical nature of '*or*', which engages decision-theoretic considerations (Merin 1999), will be crucial to predictions. The aim in this short case study cannot, of course, be to show what a boolean semantics with type shifting and with ontological hybridization cannot do. Rather, it is to show what can be done by an entirely different and rigorously algebraic semantics. This semantics is *inter alia* able to articulate into its recursive component properly pragmatic relationships that are accessible to reflective intuition about use conditions.

3 Double-Jointed Coordinate Constructions

Definition: Any instance of the sentence schema

DJCC [NP_a {*and/or*} NP_b][VP_c {*and/or*} VP_d]

is a *double-jointed coordinate construction* (DJCC) of English.

Example DJCCs are

John and Mary sang and danced.

John and Mary sang or danced.

John or Mary sang and danced.

John or Mary sang or danced.

and the 2^4 distinct instances of

{Every/Some} man {and/or} {some/every} woman sang {and/or} danced.

I offer *one* of many worked example, here for [NP *and* NP][VP *or* VP]. Parameters in MLS representations are simplified to α, β, etc.

(1) John and Mary sang or danced.

(2) John and Mary sang or John and Mary danced.

(3) John sang or danced and Mary sang or danced.

Intuitively, $(2) \simeq (1) \not\simeq (3)$. PTQ and KF standardly predict $(2) \not\Leftrightarrow (1) \Leftrightarrow (3)$. Thus, in PTQ, shorn of intensional aspects that are irrelevant for these fully extensional contexts and letting '$\|\cdot\|$' designate the interpretation function,

$$\|(1)\| = \lambda\mathcal{P}[\mathcal{P}j \wedge \mathcal{P}m](\lambda x[Sx \vee Dx]) = (Sj \vee Dj) \wedge (Sm \vee Dm) =$$
$$= \|(3)\|$$
$$\neq (Sj \wedge Sm) \vee (Dj \wedge Dm) = \lambda\mathcal{P}[\mathcal{P}j \wedge \mathcal{P}m](\lambda x[Sx]) \vee \lambda\mathcal{P}[\mathcal{P}j \wedge \mathcal{P}m](\lambda x[Dx]) =$$
$$= \|(2)\|.$$

Similarly in KF, where VPs 'sang' and 'danced' denote homomorphisms f_s and f_d from the domain of individuals I_j, I_m, etc., which are conceived of as ultrafilters, i.e. maximal consistent sets of properties, to the set of truth-values:

$$\|(1)\| = (f_s \vee f_d)(I_j \cap I_m) = (f_s(I_j) \vee f_d(I_j)) \wedge (f_s(I_m) \vee f_d(I_m)) =$$
$$= \|(3)\|$$
$$\neq f_s(I_j \cap I_m) \vee f_d(I_j \cap I_m)$$
$$= \|(2)\|.$$

KF (1985) will refer to (2) as exhibiting a 'collective' reading of *John and Mary*, while (3) exhibits the 'distributive' reading. In some cases, "pragmatic considerations" are said to select the appropriate reading. The KF example for such a case comes with the added instruction to suppose that its protagonists are a happily married couple:

(4) Every evening George and Martha go to the movies or visit friends.

Sentence (4) is said to read 'collectively', and not 'distributively' as KF would require on semantic grounds. But the 'collective reading' (2) of (1) lacks a 'pragmatic' motivation for collectivity of the kind which KF exemplify with (4). It is physically possible for each of John and Mary to dance or sing individually; they are not said to be tangoing; and there is no indication that they are happily married or cerebrally Siamese twins.

Link (1983) accounts for lexically triggered collective readings by postulating lattice-algebraic structure on the domain of individuals. Under this approach, '*and*' uniformly denotes semilattice join, $\sqcup$, in such NP-*and*-NP environments. Thus, $\|\text{John and Mary}\| = \|\text{John}\| \sqcup \|\text{Mary}\|$. Distributive interpretability, where intuitively required, is stipulated by a meaning postulate associated with a distributing VP. There will be no such postulate, then, for the VP of

(5) John and Mary carried the piano upstairs.

Link's answer for the failure of (1) to paraphrase distributively as (3) is likewise simple and direct. A VP '*sing or dance*' is not a lattice-homomorphism, but only an (upper) semilattice homomorphism, since it distributes over '*or*'. But more so than for the semilattice-union hypothesis of individual-conjoining '*and*', simplicity and directness is here the directness of brute force. The intuited reading is predicted by a stipulation which is simply a paraphrase of it, without any detours through a semilattice structure on individuals. Indeed, if '*and*' denotes join, what semilattice operation of like ontological type could '*or*' sensibly stand for? Meet is clearly not a serious candidate, and no others are in sight.

In MLS, let '*John*' and '*Mary*' for now each denote in $\mathrm{Hom}_{\mathbf{R}}([\![\mathrm{VP}]\!], [\![\mathrm{S}]\!])$). We obtain a representation which, under the standard notation conventions for linear algebra, remains close to the function-argument notation familiar from logical formalisms such as Frege's predicate logic, PTQ, or KF. This familiar notational background makes it easy to see how the laws of linear algebra yield equivalences which are unlike those of logical formalisms, and in line with meaning intuitions:[7]

(1.LH) $[\![\text{John and Mary sang or danced}]\!] = (\mathbf{J} + \mathbf{M})[\alpha\mathbf{s} + (1 - \alpha)\mathbf{d}]_{\alpha\in\{0,1\}} =$

$$= \mathbf{J}\alpha\mathbf{s} + \mathbf{J}(1 - \alpha)\mathbf{d} + \mathbf{M}\alpha\mathbf{s} + \mathbf{M}(1 - \alpha)\mathbf{d}$$

$$= \alpha(\mathbf{J} + \mathbf{M})\mathbf{s} + (1 - \alpha)(\mathbf{J} + \mathbf{M})\mathbf{d} =$$

$$= [\![\text{John and Mary sang or John and Mary danced}]\!] =$$

$$= [\![(2.\text{LH})]\!].$$

By contrast, the linear representation of (3) yields an expansion which is inequivalent to that of (2) and thus of (1):

(3.LH) $[\![\text{John sang or danced and Mary sang or danced}]\!] =$

$$= \mathbf{J}[\alpha\mathbf{s} + (1 - \alpha)\mathbf{d}] + \mathbf{M}[\beta\mathbf{s} + (1 - \beta)\mathbf{d}].$$

The predictions in each case hinge essentially on the linear algebraic properties of each of the NP and VP denotata. The constellation can also be restated more abstractly in terms of linear valuations to lead on to the alternative, tensor product representation of inter-categorial composition. For any linear form that maps the sentence denotatum to R there will be an equivalent bilinear form (i.e. a map to R linear in each of its two arguments) having for arguments NP and VP denotata. Thus, trans-categorial composition is multi-linear, here specifically bilinear. Exactly the same structure of equivalences is predicted under the tensor product explication of non-coordinative inter-categorial composition, whose presently outstanding feature is indeed the multilinearity of tensor product:

(1.LT) $[\![\text{John and Mary sang or danced}]\!] = [\mathbf{J} + \mathbf{M}] \otimes [\alpha\mathbf{s} + (1-\alpha)\mathbf{d}]_{\alpha\in\{0,1\}} =$

$= \mathbf{J} \otimes \alpha\mathbf{s} + \mathbf{J} \otimes (1-\alpha)\mathbf{d} + \mathbf{M} \otimes \alpha\mathbf{s} + \mathbf{M} \otimes (1-\alpha)\mathbf{d}$

$= \alpha[\mathbf{J} + \mathbf{M}] \otimes \mathbf{s} + (1-\alpha)[\mathbf{J} + \mathbf{M}] \otimes \mathbf{d} =$

$= [\![\text{John and Mary sang or John and Mary danced}]\!] =$

$= [\![\text{(2.LT)}]\!]$.

By contrast,

(3.LT) $[\![\text{John sang or danced and Mary sang or danced}]\!] =$

$= \mathbf{J} \otimes [\alpha\mathbf{s} + (1-\alpha)\mathbf{d}] + \mathbf{M} \otimes [\beta\mathbf{s} + (1-\beta)\mathbf{d}]$

With reference to the two variants of composition, let us specify the inequivalence of (2) and (3) in MLS. Whenever there is '*or*'-coordination, the basic MLS-representation of a sentence is a set of vectors, one vector for each disjunct.[8] The elements of this set, call them 'realizations' for the nonce, are actualized at points w of reference. Now let the relation symbol '$=_w$' stand for identity of realization vector objects in world state w. Then $[\![(2)]\!] =_w [\![(3)]\!]$ is true for any world state w such that $\alpha = \beta$. For our examples, such a world w will most likely be described as one in which John and Mary choose to perform the same kind of activity. As a fallback, w may be described as a world in which Nature has them act as if they so choose. This allegedly collective reading is nonetheless unavailable for (3), because no constraint is in place to rule out the realization of possibilities validating $\alpha \neq \beta$. The inequivalence in general of (2) and (3), which corresponds to intuition, is due to the use-conditions of '*or*', which admit as possibilities worlds w' for which $[\![(2)]\!] \neq_{w'} [\![(3)]\!]$. We cannot simply assume that $\alpha = \beta$, since we cannot assume that John and Mary will make the same choice or that Nature will make them act alike. Our pre-theoretical intuitions of paraphrastic equivalence are in conformity with this consideration: either motivated by it or by co-incidence.

Under certain circumstances, however, an assumption of relevant unanimity should be warranted on pragmatic grounds. The KF supposition for (4) is obviously intended to induce a context in which George and Martha choose the same type of activity. There are excellent rationales for them doing so,[9] but the empirical question of whether they account for the paraphrase is not thereby decided. Indeed, the answer is negative. The rationale appears not to be necessary, since (1) reads analogously to the allegedly 'pragmatic' reading of (3). Neither is it sufficient, as is seen from comparing

(4′) Every evening George and Martha went to the movies or visited friends.

(4′F) Every evening, George went to the movies or visited friends and Martha went to the movies or visited friends.[10]

It is clear that our assumptions about happy marriages do not suffice to force for the nominally less informative (4′F) a reading equivalent to that of more informative (4′).[11] Here one should appeal to 19th century common-sense to say that non-use of the shorter and more informative (4′) ensures that coincidence of ($4'F$) with (4′) is all but ruled out. Give or take the adsentential prefix, '*every evening*', a pragmatics of language use rationalizes the resistance of selection-functional freedom of '*or*' to being constrained by pragmatic world-knowledge type assumptions about the world-coordinate of the context.

4 Convex Coordination: Non-Distributing '*and*'

'Collective' is a highly general term as commonly used. Applied to intuited readings of (1) or (4), it referred to unanimous choice of the same action type by a plurality of agents. However, this did not rule out that each agent did individually perform an action of the common type. Thus, while sentence (1) expands to (2), that expression will in its turn expand to

(2E) John sang and Mary sang, or John danced and Mary danced.

The identity $[\![(2)]\!] = [\![(2E)]\!]$ shows that '*and*' in (1) and in (2) is 'distributing' after all. If a rationalization of the identity $[\![(1)]\!] = [\![(2)]\!]$ is wanted on top of multilinearity of NP·VP composition, it will derive from the properties of '*or*'. Specifically, it should imply an equi-choice constraint on the two selection functions associated with strictly less informative (3). Properly collective '*and*' appears in sentences in which '*or*' need not occur, as in

(5) John and Mary carried the piano upstairs.

This sentence is fairly unambiguous if one knows about people and pianos. The MLS account will reconstruct it, now just in tensor product format, by

(5.LT) $[\![$John and Mary carried the piano upstairs$]\!] =$

$$= [\alpha\mathbf{J} + (1-\alpha)\mathbf{M}] \otimes \mathbf{ctpu} =$$

$$= \alpha\mathbf{J} \otimes \mathbf{ctpu} + (1-\alpha)\mathbf{M} \otimes \mathbf{ctpu} \quad (0 < \alpha < 1)$$

$$\neq \mathbf{J} \otimes \mathbf{ctpu} + \mathbf{M} \otimes \mathbf{ctpu} = [\![\text{John carried the piano upstairs and Mary carried the piano upstairs}]\!].$$

In geometry, convex combination explicates the fundamental concept of 'betweenness'. A point **b**, considered as a point vector, is between points **a** and **c** considered likewise iff it is a convex linear combination of **a** and **c**.[12] The felicity of the MLS representation to English-speaking minds is attested by a securely disambiguating paraphrase of (5), '*John and Mary between them carried the piano upstairs*'.[13] The relevant instance of convexity is not the physical way pianos are usually carried, i.e. slung between people rather than placed atop their heads. This is seen from the adnominal position of the infix and from '*Kim and Sandy* (*between them*) *owe Lee five pounds*'. Economic and juridical reasoning rationalize convex combination. Neither John nor Mary are physically fractioned, but credit or responsibility for the event in which they participate as agents is shared out among them. Frege's puzzling dictum "an expression has meaning only in the context of the sentence", was never truer than here; and it gets its situated meaning from the relevant space of ultimate values. Agents John and Mary are considered less with regard to their physical or referential being than as authors, *sub specie utilitate*.[14]

A somewhat similar example which does not admit '*between them*' as a disambiguator is the title of a 1980s British motion picture:[15]

(6) Sammy and Rosie get laid.

By default, (6) will be read as describing reciprocal action.[16] The plot, however, sees each of Sammy and Rosie in amorous embrace with sundry others, though never with each other. The expectations generated by the title are thus disappointed by the script,[17] which validates a distributing reading of '*and*':

(7) Sammy gets laid and Rosie gets laid.

The MLS representation for the non-distributing title default is again by way of proper convex combination. It should thus capture the reflective intuitions (i) that Sammy and Rosie jointly manage to bring about one event of getting laid, be it occasional or habitual, and (ii) that neither can claim prioritized entrepreneurial authorship for it.[18]

Convex combination as deployed in MLS models an everyday relationship that might be called '*entanglement on the cheap*' because the scalar coefficients on '*and*'-summands stand in a non-degenerate relation of functional dependence. Their exact values in applications such as (5) or (6) are generally unknown and in this they differ from the constant unit coefficients of clause-conjoining '*and*'. Nonetheless, whatever value one of the coefficients might take will always determine the value of the other.[19] The situation differs from that of (7) in which Sammy and Rosie are not so entangled. They each go off under their own steam and have amorous encounters with partners that remain anonymous for the purposes of (7). Each of these events of getting laid may well have the responsibility for (or the value created by) its performance shared out between Sammy and some partner and between Rosie and some partner, who are all known as characters to the movie audience. Yet the feel of

(7) considered in its own right makes each of Sammy and Rosie appear somewhat entrepreneurial in realization of the respective events—at any rate by comparison to their openly mutual participation in (6).

The reading for the subject noun phrases of (5) and (6) which is prescribed by the MLS treatment is structurally akin to that for the adjectival phrase in

(8) This newspaper page is black and white.

Sentence (8) does not license the inference to either of '*This newspaper page is black*' or '*This newspaper page is white*' and it was in all essentials fielded by Chomsky (1955:40) as a counterexample to logical semantics.[20] That a rendering by convex combination is appropriate for (8) is even more obvious than for (6). The page surface, after all, is covered by white and black areas in proportions that are greater than 0 and less than 1 for each, and that must sum to 1 by definition of proportion.

The '*and*' of exx. (5), (6) and (8) or their readings cannot readily be subsumed under the label 'collective', because it is not obvious that it fits '*black and white*'. If anything the latter designates a mixture, while '*Sammy and Rosie*' or '*John and Mary*' do not intuitively stand for mixtures of individuals. 'Non-distributing' would do, because the examples satisfy a properly empirical constraint on what MLS identifies as the convex combination reading of '*and*'. The reading is available only for subsententially phrasal coordination. Its unavailability for clause- or sentence-coordinating '*and*' guarantees the intuitable import which is manifest in a failure to distribute. On the other hand, 'non-distributing' is unhelpfully non-specific, since it is frequently applied to DJCC contexts and might also describe the failure of Quantum Logic's explicatum of '*and*' to distribute over its explicatum of '*or*' (Section 5.2 below). Labelling a 'non-event' type phenomenon with one's pet positive theoretical description is a lesser sin than ambiguity, and so I propose that we call phrases exhibiting this reading of '*and*', *convex coordinate constructions*.[21]

Note four advantages of the analysis of the whole range of exx, (1)–(8). First, there is no need to undo predictions of a basic semantic framework. Neither brute force nor a mid-clausal change in the underlying formal ontology[22] is required. Secondly, the common structure of NP-coordinating and VP- or AP-coordinating '*and*' in (5) and (8) is brought out. (5) illustrates an economic value interpretation of linear spaces, while (8) instantiates the 'image' or physical type. Thirdly, the common structure rests on substantive interpretability by sharing or mixture. Approaches by way of semilattice join and a meaning postulate for distributing contexts are neutral in such respects, which may not be an advantage. Fourthly, and most generally, recall that scalars as they appear in linear combinations are technically speaking operators on the additive group of the linear space, which is the set of all vector sums and differences. These scalars may be values of functions which take for arguments components of the context of use, including co-text. They are thus available to represent agent or random choices and to code up contextually and lexically induced constraints on interpretation, including e.g. 'respectively' constructions (Merin 1988, 2002). Nothing like this is available in boolean and more generally in lattice algebras. Hence, any provision of functionally similar facilities would be extrinsic to the core structure of coordinative (and thence quantificational) semantic recursion.

The applicative type assignment to NPs, exemplified in (1.LH) was typed so as to have the NP denote a map from VP denotata to sentence denotata. This was for expository simplicity. On the linear algebra side, its expository advantage is that it employs applicative structure and notation which is familiar from early undergraduate mathematics. On the side of language studies, it also happens to shadow the applicative structure of NPs and VPs in Montague semantics and modulo function/argument role reversal, that of Fregean first order logic. Both applicative structures have correlates in syntactic analysis of expressions of type 'Sentence' by concatenation of constituent typed expressions, many of which are obtained by algebraic residuation. However, nothing essential to predictions will change if we reverse the type assignment of NP and VP, nor if we substitute a tensor product interpretation in which neither of NP and VP denotes a semantic argument to the other.

The multilinearity of tensor product leads to the same predictions in respect of distribution behaviour as does the applicative route.[23] The choice of route, be it applicative or by tensor product, will thus have to be made (i) by reference to theories of open class word content relations and also (ii) so as to minimize appeal to gratuitous assumptions. The tensor product route appears to be advantageous on both counts. Tensor product in the category of real vector spaces can serves as a substrate for representing just about any association the computer scientist wants represented. For present purposes, accordingly, what affords its intrinsic predictive purchase is its multilinearity together with the assumption that coordinate phrases and clauses denote linear combinations in algebras over orderable and thus infinite rings or fields.

5 Intratheoretical and Intertheoretical Context

The first three sections above are, save for a few stylistic expansions, the contents of a short paper submitted to an annual conference venue in 2006. Like nearly all my earlier ventures aimed at making work on linear semantics accessible to the public with someone's reassuring certificate of goodness on it, this one proved uninspiring. The appearance in print of Preller & Sadrzadeh (2011) suggested to me that perhaps now there is a public which has both (i) an interest in languages spoken by people outside of a regimented context and (ii) the mathematical culture which readies the mind for the exploration of alternatives to overworked approaches to the domain of (i).

In this section, I first place the small focal case study of Section 3 and its companion of Section 4 in the context of the larger research project from which it is extracted. Moving outward from the language fragment in effect considered, there are some hints on the treatment of negation. Relations to the well-motivated doctrine that meaning is, or is above all, truth conditions will also be addressed. MLS is not truth-conditional *per se*. It is not even proof-conditional, as 'anti-realist' semantics by way of intuitionistic logics are commonly conceived of. In the second subsection, I relate MLS briefly to recent work of which Preller & Sadrzadeh (2011) is both a summary record and an extension. To keep things in proportion to present purposes and so as not to pad the bibliography with hearsay, I shall treat their article both as a sufficient statistic and as a bibliographical repository of this body of work.

5.1 Intratheoreticals

A referee for the conference venue to which the case study was submitted in 2006 agreed that, yes, the MLS approach did predict the above data, where familiar approaches did not. However, so (s)he also noted, no account was given of why the preference for the collective reading disappears e.g. in downward monotonic contexts such as '*Every time that John and Mary sing or dance the neighbours become nervous*'. Here John singing and Mary dancing will intuitively suffice to make neighbours nervous. Both of them singing or both of them dancing, which is a special case, will do so too, but is not presented as being a requirement. Familiar logical approaches, so the referee's claim was, did entail that there was no preference for the collective reading here. It was not apparent, by contrast, how the MLS proposal would so predict.[24] Hence, even though it predicted correctly for DJCCs in the usual assertoric upward monotonic contexts, while boolean/logical approaches failed, it had no overall advantage over these standard approaches, and so should not be presented.[25]

What went unobjected-to was an assumption that is shared by many students of natural language in the wake of transformational and early North-American structuralist grammar. This is that the principal criterion for the descriptive viability of a theory of meaning is its ability to account for judgments of paraphrastic equivalence. This is an eminently algebraic criterion, since the axiomatizations of all familiar algebras are predicated on the concept of equality and non-equality of terms.

By contrast, already a notion such as entailment, $A \models B$, which is not restricted to the special case of equivalence, does not have equally intuitive algebraic analogues, *vide* $A \wedge B = A$ or $A \vee B = B$. In boolean algebras, the 'natural' partial order structure on the algebra explicates entailment relations. Similarly, there is a natural, intrinsic lattice structure of subspaces in any linear algebra and it can be utilized to similar purposes. But notions which are intuitively more sophisticated than equality of objects, notions such as that of an 'ideal', are required to explicate entailment algebraically. Only in this theoretical context do solution sets to equations such as those in which arise in quotienting or residuation acquire their intuitive meaning in terms of entailment and related implication operations.

But entailment in one of these senses is not the only option to consider. We could also think of orderings among elements of real vector spaces, orderings that are not simply predicated on the subspace lattice structure of, say, $\mathbb{R}^n$. Thus, any linear functional on $\mathbb{R}^n$, induces a total (though not in general antisymmetric) ordering on the vector elements of the space. The functional's value space is $\mathbb{R}$, which is totally ordered, and the ordering on elements of $\mathbb{R}^n$ will be by membership in pre-image equivalence classes. The same holds for modules over $\mathbb{Z}$, in which all scalar coefficients are integers.

The hypothesis for MLS is that use-valuations are linear functionals on the vector spaces of sentence denotata. Our ordering criteria in empirical or economic applications might thus be determined by preference relations. It is an empirical question, then, whether such orderings will predict intuitions of consequence which are not already intuitions of equivalence. Under the very general assumption that meaning is the imposition of constraints on expectations (be they credibilities, desirabilities, or credibility weighted mixtures of the latter) and on possible actions, vector elements in a relevant subspace which do not satisfy a value constraint will be ruled out as admissible by an utterance. On this basis, a contextual entailment-like relation will be induced, which is truth-based like all equational constraints on any kind of space.

There is precedent for a preference-based approach to reasoning in a more traditional semantic framework with propositional denotata. 'Argumentative scales' a.k.a. 'pragmatic scales' have generated a small industry in linguistics. The mathematical approach to them taken in Merin (1999) demonstrates that certain prominent phenomenological order relations which have been conceived of as being generated by entailment relations are not, in fact, generated by them. Instead they are generated at base by relations of preference or, as a special case, by ordinal relations of signed evidential relevance as explicated in the probability calculus. This precedent is in a domain where preference properly competes, but also co-operates with entailment as an explicandum of inferential intuitions. It should thus make somewhat less surprising a related approach in a more unfamiliar formal ontology such as that of real linear spaces.

The theory of negation in MLS makes essential use of prefererence-based order structures. To someone familiar with negation in classical logic (or Heyting lattice pseudocomplements or subspace lattice orthocomplements) this will seem a little strange. After all, negation and any one of conjunction, disjunction, and material implication provide a functionally complete set of sentential connectives for classical, truth-functional languages and by extension their predicate-logical and intensional extensions. On the other hand, a philosophical logician who tries to make descriptively adequate sense of negation in natural languages will know that truth-functional recursion leads to many apparent dead ends for which no breakthrough is in sight. A familiar example is the non-truth-functional interaction of negation with ordinary language conditionals. There are less well known elementary surprises too, such as the widespread failure of German to conform to De Morgan's Laws, of which even substructural logics such as Linear Logic have correlates.

The upshot of these unsolved problems must be, at the very least, that negation in natural languages is not as well-behaved as the boolean complement happily is. Negation in MLS is treated more fully and most recently in my booklength treatment from 2002. A short treatment of one of the general principles is accessible in the section entitled 'De Morgan intuitions without binary booleans' im Merin (1994).

Briefly summarized, the effects intuited for negation depend on speech act type, say, on whether an utterance to be denied is a claim (demand, command, assertion) or a concession (permission, admission). What distinguishes these two broad act types are *inter alia* inverse preference relations relative to an intersubjective default description of preference objects. As economic beings and as a matter of brute definitional fact we prefer *ceteris paribus* to maximize claims and to minimize concessions. In a linear semantics, these acts in turn generate 'cones' or *half-spaces* of vectors verifying a constraint. For a claim, this will be the set of all vectors in a subspace determined by atomic subclause denotata[26] which are at least as highly valued in the context as the speech-act-independent sentence-denotatatum, call it v_σ. A representative element chosen from this half-space of the subspace will be the object of the claim. For a concession, '*at least*' is replaced by '*at most*'. The negation of the sentence expressing v_σ is by default interpreted by an act of denial of a claim for v_σ. It picks out the set of vectors, within the same subspace, which are valued *less than* v_σ, each time with respect to a given canonical value-orientation. For denials of concessions, the dual constraint description, i.e. '*more than*' holds.[27] Negation at its most general will thus be *half-space formation* within a subspace, followed as always by selection of a half-space element. In the description language of Merin (1988,2002), this vector object will be given as a Hilbert ϵ-term picked from the cone by choice function semantics (Asser 1957). This recovers the familiar association of negation as complement-formation of one kind or another. By linearity of linear functionals, additive-inverse formation, which explicates certain pre-theoretical and largely extra-logical notions of contrariness, will be a special case of constraint satisfaction.

The issue of negation once more raises the general issue of truth conditions. How can truth conditions be determined when they are not operationalized as paraphrase conditions? Rather complex experimental setups may be needed to elicit judgments when paraphrastic means are unavailable. They might involve exposure to pictures or to toy model installations, or to a life-like experience. This brings me to a very general point about MLS, both negatively and positively.

MLS vector objects and sets of such objects *do not specify truth conditions*, as propositions or proposition-denoting sentences do by definition. Here is a toy example of the contrast: $A \wedge B$ where A and B are propositions, is itself a proposition and thereby has truth conditions just as A and B have. By contrast, the vector object $2 + 3$ from the module $\mathbb{Z}$ of integers has no truth conditions. What does have truth conditions is, for instance, the equality

$2+3=4+1$ or the inequality $2<1$. Each is a proposition. The first is true, the second is false. Each of $=$ and $<$ is a two-place predicate that takes numbers or other orderable entities for arguments and has a truth value for value. Similarly a linear equation or a system of linear equations is a proposition (and, when brought into homogeneous form, determines a subspace of the given space).

The toy example of the distinction also comes with a familiar toy fact that cuts right across term algebras. $2+3=4+1$, which might be a statement in the theory of $\mathbb{Z}$, has the same truth conditions as $\sqrt{9}=3$, which is a statement in the theory of $\mathbb{R}$, and as $\neg(A \wedge \neg A)$, which is intuitively a statement about logic. Likewise, $2<1$ has the same truth conditions as $A \wedge \neg A$. The toy fact will not daunt the practitioner who supplements truth-conditions by verification conditions and perhaps syntactic recursion-conditions. All these will be variations, one hopes interesting variations, on Carnap's idea of "intensional isomorphism". But even so the example will suffice to show one thing. It is an act of great faith to maintain, as many do, that truth conditions capture the essential meaning of 'meaning'.[28] That claim may yet be true, but to be at all impressive, it should not eventually pan out to "the meaning of a sentence σ is given by the truth conditions of nearly everything that can truly be said about σ".

Are truth conditions ubiquitous in natural languages? Imperatives do not have truth conditions. When someone addressess one of them to you, you cannot reply: '*That's true*' or '*That's false*'. The ubiquitarian reply will be that imperatives have satisfaction conditions with like structure. You can reply '*No*' and '*Yes*' to them. What they specify is what must be, or what ought to be or what may be. Which of these it will be depends depend on whether they are read as commands, suggestions or permissions. Indicatives specify what is or was or will be. Just as we can check whether something is the case, so we can check whether an imperatival command or demand has been complied with or, in the case of concessions and permissions, taken up. Questions, in their turn, can be handled partly by way of the indicatival or imperatival answers which they admit of, and partly by way of the imperatival element in soliciting some answer. Exclamatives, under which I subsume interjections such as '*Oops!*', are a harder nut to crack. To identify their meaning, as David Kaplan has suggested, with that of a description of their context of felicitous utterance strikes me as being just too indirect and missing out on their peculiarities.[29]

But suppose we are restricting ourselves to simple indicatives. And suppose we consider once more our toy examples of truth-conditionally equivalent sentences from mathematics that have little or no apparent subject matter in common. These examples suggest that there is scope for representations of meaning that need not themselves have truth conditions, but could serve as a substrate for truth conditions. The latter, in their turn, might or might not be fully compositional.

The reply to this consideration from the defender of purely truth-conditional semantics could be: Why not simply let the syntactic sentence string, parsed into words or morphemes and then upwards into phrases and clauses be that representation? This suggestion resembles Norbert Hornstein's quip that the semantic objects of Jon Barwise's and John Perry's 'Situation Semantics' were simply "de-frocked sentences".

Our response to such an appeal to Occam's Razor will be: Yes, our judgments of paraphrastic equivalence may be in accordance with rules of transformation in the sense of Carnap (and early Chomsky and ultimately Hilbert) which are not predicted by the truth-conditional, would-be compositional interpretation. However, we cannot just say that they are rules of transformation of some formal syntactic calculus that simply fail to conform to the predictions of a truth conditional interpretation. Such a claim will intimate, and will derive all its interest from the assumption, that these rules are in crucial respects arbitrary. But this assumption cannot be sustained when the observed paraphrastic transformations behave as they do, for one, in MLS. The MLS rules conform to the laws of linear algebra. And these laws have a wide range of independently motivated models in the world—enough so for us to speak of a linear semantics without having to insert a hedge 'technically'.

In what general relationship would MLS representations stand to truth conditions? Surely, equal truth conditions should attach to MLS-equivalent sentences, if they attach.[30] However, the basis on which MLS denotata of compound sentences are computed will not in general be by recursion on truth conditions of constituent sentences. Analogous strictures will then hold for sub-sentential recursion of meanings. MLS denotata are in this representational respect like pictures or other images.

Pictures do not generally have truth conditions in any direct way.[31] When they acquire them, they have them in ways for which compositionality may be hard to claim, and if at all, then with a restricted expressive range. They might indeed be already predicated on prior interpretation of the picture by means of a language of sentences that have truth conditions. This sounds like an argument for claiming that an MLS for a natural language will be entirely redundant, because we cannot reasonably do without a truth-conditional or similar proof-conditional semantics, while we might well do without a use- or relevance-conditional semantics. The reply will be twofold. First, an R-linear semantics appears to inform our intuitions of acceptability and paraphrastic equivalence as a matter of brute empirical fact. Secondly, its representations need be no more redundant to us than pictorial images that have acquired truth conditions or than imagistic thought. Philosophers at the very least from Aristotle onwards have suspected imagery to be indispensable to unaided human reasoning.

How an MLS-interpreted language acquires its truth-conditional import will, I think, have to be explained by the use made of it in arguments. It is arguments alone which can properly be said to have logical forms. This is surely the message of the method of 'disambiguated languages' that is employed in Montague semantics. When logical form is taken, as it is in applied classical logic, to specify truth conditions, it is arguments and their clause-level constituents which have truth conditions. Sentences will have them only mediately so, after a manner of speaking.

5.2 Intertheoreticals

Recall that one of the two distributive laws which are predicted to hold by boolean or classical semantics fails empirically in everyday natural language (NL).[32] Reported this way, the phenomenon might raise hopes that quantum logic (QL) as formulated by Birkhoff & von Neumann (1936) for experimental observations in quantum mechanics (QM) is appropriate for NL. Such hopes cannot be sustained on this particular count. The usual logico-linguistic associations of QL are (i) of '*and*' with meet (minimum on familiar 0-1 truthvalue pairs) and (ii) of '*or*' with join (maximum on familiar 0-1 truthvalue pairs) when each is a *bona fide* sentential connective. The association will have to hold *a fortiori* when both words have occurrences in the same sentence. Granted this translation lore, it is the wrong law which fails in NL. Everyday natural language has no problem with that half of the pair of dual distributive laws which fails in the observation part of QM and thence in QL.

Example: '*Kim walks and either Sandy talks or Lee sings*' is intuitively equivalent to '*Kim walks and Sandy talks or Kim walks and Lee sings*'.[33] Under the canonical translation lore, this equivalence is an instance of this law of distributive lattices:

$$\vdash A \sqcap (B \sqcup C) = (A \sqcap B) \sqcup (A \sqcap C).$$

But this is the law which fails in QL, and the intended interpretation in QM affords no loophole by dualization. By contrast, what fails in NL is the equivalence of '*Either Kim walks or Sandy talks and Lee sings*' with '*Kim walks or Sandy talks and Kim walks or Lee sings*'. The corresponding lattice-theoretic law, dual to the first, is

$$\vdash A \sqcup (B \sqcap C) = (A \sqcup B) \sqcap (A \sqcup C).$$

This law is validated by QL under the usual translation. The distributive lattice laws cannot be derived from the absorption laws, which are validated by any lattice. The absorption laws, in turn, cannot be derived from the idempotent laws, which are respectively validated by any join and meet semilattice.

By contrast, MLS predicts the reflective, semi-theoretical intuition that the failing halves of the three pairs of laws are all reducible to the presence of a certain interpretive possibility. Vector objects, if interpreted by linear functionals in terms of evidential values, would violate the idempotent law that governs the classical norms of evidence.[34] Presenting the same evidence twice, so the rationale for the classical rule, ought not to strengthen one's argument. Additivity of relevance, to a hypothesis H, of the relevances of boolean conjuncts is feasible for suitable signed relevance measures defined in terms of probability measures P on the proposition algebra, in particular the log-likelihood ratio measure. Additivity is guaranteed when conjuncts are independent conditionally on the cells of the hypothesis partition $\{H, \overline{H}\}$. Such conditional independence is a candidate for a default presumption for assertions made to a point (Merin 1999). If relevance is our utility, then **R**-linear additive '*and*' serves us well. However, a proposition A that is at all relevant to at least one proposition in an evidential context given by a probability measure P can never be conditionally P-independent of itself. The only predicament where relevance additivity does not fail for the schema '*A and A*' is when A has zero relevance to all propositions B under P. But this case is excluded by the presumption that the utterance is relevant.[35] The role of classical probability measure, a real-valued set-function, in relating boolean and linear algebras is obvious enough, and well before we consider extensions of the probability measure concept on linear subspace algebras as in QM. Evidential relevance defined in terms of probability is one link between linear and truthconditional representations of natural language sentences that jointly determine intuited meaning.

In 1988 Martin Hyland of Cambridge University pointed out to me the rather abstract relationship between Linear Logic [LL], which had recently been formulated by Jean-Yves Girard, and linear algebras, indeed via an abstract notion of 'game'.[36] Game semantics for linear logics appear to be the most natural semantics available for them, whereas classical and intuitionistic logics have established alternatives that can hold their own against game interpretations. LL is resource-conscious in using up premisses, but affords implementations for intuitionistic and classical logic by means of resource-expanding operators. Though it does have a non-idempotent conjunction-like connective and others with suggestive choice-of-resource interpretations, I did not manage to obtain a reconstruction of the data on its basis.

Recent work of Preller & Sadrzadeh (2011) offers semantic models for first-order-logic-like fragments of natural languages in linear spaces, which are based on a very abstract consideration. The categories of (i) finite-dimensional vector spaces, (ii) of sets and 2-sorted functions with sorts for sets and for their elements (as instantiated in two-sorted first-order logic) are each symmetric compact 2-categories. So is (iv) quantum logic [QL], which has models in the lattice of projectors in finite-dimensional Hilbert spaces and has recently been applied to information retrieval (van Rijsbergen 2004, Widdows 2004).[37]

Preller & Sadrzadeh show that the compact 2-category of finite dimensional semi-modules (a generalization of finite-dimensional linear spaces) over a bounded lattice of real numbers (say, on the unit interval) afford representations for both the QL of semantic vector models and the functional models. The explicit motivation is to *represent a truth-conditional semantics in vector space models*.[38] The most intuitive way for the general reader to begin appreciating the connection is to note that the subspaces of a vector space (as distinct from its vector elements) form a lattice. For example, in Euclidean 3-space, two planes that each contain the origin as a point are subspaces of 3-space, their meet is a line, another subspace. The orthocomplement of a subspace of an inner product space is the set of vectors whose inner product with the elements of the subspace is zero.

The subspace lattice is not boolean, but under suitable conditions on vector spaces it will have a non-degenerate boolean sublattice. Even without full boolean structure, the standard lattice operations are generalizations of conjunction and disjunction. A negation-like operation is available in the concept of an orthocomplement, and so is a residuation operation with core properties of implication.

MLS in its basic form does not exploit the subspace or projector algebra for the formation of connective and quantifier interpretations. It exploits subspaces much as one does in boolean semantics and their pragmatic extensions when looking for smallest algebras that nonvacuously satisfy a set of constraints. More significantly, though, some of its predictions of unacceptability derive from the interplay of pure MLS with informational-evidential constraints that reflect in the idempotence of lattice meet, which instantiates to classical conjunction. Thus, '*Kim is blond or Sandy talks and Kim is blond or Lee sings*' is analyzed in MLS as envisaging a possible realization, out of four posssibles, which is equivalent to '*Kim is blond and Kim is blond*', i.e. the vector sum of the two conjunct denotatations (Merin 1997, 2002).[39]

There is a mismatch between the linearly predicted two-fold evidential value and the facts of evidential norms predicated on the probability truism

$P(A \wedge A) = P(A)$. The claim is that this mismatch is responsible for the inacceptability of '*Kim is blond or Sandy talks and Kim is blond or Lee sings*'.[40] The same kind of reason explains the empirical failure of just one of the lattice absorption laws. The mappings to linear algebras over the finite, two-element ring or field $\mathbb{Z}_2$ a.k.a. $GF(2)$ (Preller & Sardrzadeh 2011) in terms of whose operations the familiar boolean operations are definable, afford idempotent structure for '*and*'. Structure of this kind, however we obtain it, is required to lead to the predictive mismatch with linear additivity.

6 Conclusion

I take the class of examples just examined to be paradigmatic for a *Dual Representation Thesis*. The thesis says that a linear semantics in the sense envisaged and a truth-conditional semantics jointly and interactively generate our basic intuitions of compositional meaning. The truth-conditional component will be acting as a check on pure MLS-representations without necessarily affording full compositionality by itself. (Recall the example of conditionals and negation.)

A wider systematic and historical perspective will associate the two principal descriptive components, non-truthconditionally linear and truth-conditionally logical, with two classes of the traditional Kantian *a priori* of pure reason. The logical, finite ring component of the semantics would correspond to *analytic a priori* constraints on colloquial human throught. The pure linear component of MLS would correspond to *synthetic a priori* constraints, which Kant associates with the laws of quantity and in particular with the pure intuitions of space and time. To this we might add parts of the metaphysics of practical reason considered at its broadest so as to encompass the micro-economic realm that Kant would refer to as 'pragmatic'.

Kant's doctrine of pure reason is inadequate as a would-be transcendental theory of the necessary conditions for the possibility of experience of any intelligent being, This much has been shown by the discovery of alternative logics and geometries, some of which have been privileged by 20th century physics over Kant's choice. However, as an outline doctrine about the workings of human minds on their best everyday behaviour, it may still extend conservatively so as to approximate some core facts of cognitive social psychology.

Natural language, whatever transcendental ambitions some philosophers have for it, is an empirical phenomenon. Even the norms which it works by are empirical, sociological, or biological facts once we describe them, as distinct from acting within them or on them. It would be a philosophical category mistake to think that we are not moving in the realm of theoretical and experimental social psychology when we prioritize descriptive adequacy over superior information engineering performance. This makes descriptive formal semantics of natural languages a branch of mathematical psychology even while it remains part of philosophy. Philosophy, supposedly, aims to establish not only what ought to be, but also what is. Semantics so conceived relates to pure information engineering as distinct from computational linguistics at most through considerations of cognitive ergonomy. The same aesthetic criteria favouring the parsimony of a maximally unified framework apply here as in other branches of mathematical philosophy and mathematical linguistics.[41]. Again, the same aesthetic criteria apply as in all other scientific instances of the speculative game 'What if we replaced the x we take for granted by some y of independent interest?'. But unlike in strongly normative branches of philosophy or theoretical computer science, the emphasis will be on accounting for the phenomena rather than on shaping them in line with practical and aesthetic preferences.[42]

Notes

[1]How well and, above all, how non-trivially UG and other familiar frameworks for semantic composition do so has been disputed. See Kracht (2007) and the literature cited there. We ignore this foundational problem which must affect any kind of putatively compositional semantics. — N.B.: While some background in elementary logical semantics, linear algebra (and, for the comparative remarks of Section 5, in category theory) is of necessity taken for granted, efforts will be made to recall salient principles for general readers and to motivate for them key points which are not readily addressed by key word search.

[2]In this case, vectors $\mathbf{v}, \mathbf{w}$ will be considered as functions from points of reference c, say contexts c of use, to $\mathbf{R}$, such that where $\mathbf{v}(c) = r_1$ and $\mathbf{w}(c) = r_2$, $[\mathbf{v} + \mathbf{w}](c) =_{df} \mathbf{v}(c) + \mathbf{w}(c) = r_1 + r_2$. The leftmost '+' symbol stands for an operation, vector addition, which is different from that denoted by the rightward '+' symbols, which will be addition of real numbers when $\mathbf{R} = \mathbb{R}$. When vectors $\mathbf{v}$, $\mathbf{w}$ are considered as functions, the operation structure of $\mathbf{R}$ is lifted to the function space, where its axiom schemata now hold for vector addition.

[3]Think of $a \in \mathbf{R}$ as a map $\mathcal{L} \rightarrow \mathcal{L}$ such that, e.g. $a : \mathbf{x} + \mathbf{y} \mapsto a(\mathbf{x} + \mathbf{y})$ equivalently $a\mathbf{x} + a\mathbf{y}$. — A linear space is closed under linear combination just as a boolean algebra is closed under boolean polynomial formation. Just as an atomic boolean algebra is the set of all boolean polynomials of its atoms or any other generator set of elements, so a linear algebra will be the set of all linear combinations of some set of basis vctors, which need not be unique.

[4]If $\mathbf{A}$ is a linear map $\mathbf{y} \mapsto \mathbf{A}\mathrm{y}$ and $\mathbf{y} = a\mathbf{w} + b\mathbf{z}$, then $\mathbf{A}\mathrm{y} = \mathbf{A}(a\mathbf{w}) + \mathbf{A}(b\mathbf{z}) = a\mathbf{A}\mathbf{w} + b\mathbf{A}\mathbf{z}$. Also $[c\mathbf{A} + d\mathbf{B}]\mathbf{x} = c\mathbf{A}\mathbf{x} + d\mathbf{B}\mathbf{x}$. A linear form on a space V is a linear map in $\mathrm{Hom}_{\mathbf{R}}(V, \mathbf{R})$.

[5]'Relatively' means that type shifting regimes in Montague and similar semantics are still fairly restrictive in the shifts they allow.

[6]This use of type shifting (TS) is orthogonal to 'thema vs. rhema' phenomena, which are today rubricised under pragmatics of discourse or 'information structure'. Formerly, scholars associated the 'rhema'/'comment' vs. 'thema'/'topic' distinction with that of function vs. argument. There is no mathematical rationale for this. For all the categorical elegance of TS, the lack of such independent motivation for TS should cast doubt on the tenet that sentences of English, as distinct from the arguments they are used in, have a logical form.

[7]The label schema (x.LH), where $x \in \mathbb{N}$, indicates the linear semantic object representing example sentence x under the $\mathrm{Hom}_{\mathbf{R}}([\![\mathrm{VP}]\!], [\![\mathrm{S}]\!])$ for the NP. Similarly, (x.LT) will label semantic objects under the tensor product representation.

[8]Features of the speech act determine whether an inclusive, exclusive, or tending-to-exclusive reading is obtained. Put briefly, claims, prior to upperbounding by implicature-like constraints, will suggest an inclusive reading, concessions an exclusive reading. See Merin (1988, 2002), and for the general principle, Merin (1994).

[9]Their tastes might conform or they each prefer doing the same thing sufficiently much to doing things alone and let a random device or calendar rule choose among visiting and movies.

[10](4′) is obtained from (4) by changing present to past tense inflection, which neutralizes the distinction between singular and plural number inflection that might induce thought-experimental artefact.

[11]The situation is akin to that for '*Kim guessed a number and Sandy guessed a number*' and even more extreme than that for '*Every Italian loves a German*' which is not normally read as equivalent to the more informative '*There is a German who is loved by every Italian*', but can just about be imagined to have such a reading. A

logical treatment of indefinite determiners '*a*' or '*some*' in terms of choice function semantics for Hilbert ϵ-terms, or a related Skolemization of existential quantifiers, will make plausible the analogy.

[12]In Hilbert's axiomatization of geometry, unlike in Tarski's, basic 'betweenness' is strict, $\mathbf{a} \neq \mathbf{b} \neq \mathbf{c}$.

[13]The non-collective, non-convex, distributing reading can be forced with '*John and Mary each carried the piano upstairs*', so there is scope for disambiguation, even if the default is strongly for the 'collective,' convex combination reading.

[14]In economic input-output analysis a production process is represented by a linear map on a space of commodity vectors. It is this analogy which might sometimes offer a substantive motivation, with implications in naive concepts of Agency, for the applicative format of transcategorial recursion.

[15]Directed by Stephen Frears and scripted by Hanif Kureishi, the movie is an exercise in moral philosophy conducted by means of audio-visual narrative semifiction.

[16]One could expand (6) by tagging on '*with each other*' to make the implied reciprocity explicit, if the expansion were not so unidiomatic. Reciprocity as distinct from mere co-participation or co-responsibility, which also reflects in the medio-passive or 'middle' voice of the verb phrase, might explain why '*between them*' cannot be infixed into (6).

[17]This element of humour presumably motivated the choice of title. The everpresent backdrop to the movie's personal subplots are the socio-economic developments of the period, which are depicted as dashing earlier hopes and offending moral sentiment. They might just warrant an additional and somewhat euphemistic reading of the title in which the passive takes over from the middle voice.

[18]I guess that '*and*' in such implicitly reciprocal or participational constructions would be a candidate for a representation by entanglement in quantum-mathematical

formalisms for lexical semantics. This non-separability concept is proposed in Melucci & van Rijsbergen (2011) for representing compound predicates such as ‘pet fish’. I adapt the term in a moment.

[19]MLS-‘*or*’ also exhibits disjunct entanglement on the cheap. Here any incidence of mixture will be that of purely epistemic probability mixture. Inclusive readings are generated, where intuitive, by means similar to those which explicate negation, on which see below.

[20]Chomsky never recovered faith in logical semantics, if any he had had. Aristotle has a structurally equivalent example of fallacious distributing inference (involving a darkskinned face and white teeth) in Ch. 5 of the *Sophistical Refutations* at 167a.

[21]The naming policy mirrors that of Keenan & Faltz (1985:270) who, in a section titled ‘Non-homomorphic Predicates?’, call a non-distributing NP-coordinating ‘*and*’ with similar phenomenological properties a “higher order ‘*and*’” which “(roughly) forms sets”. (Set-formation is a predictively equivalent to plural-individual formation by semilaticie join.) Merin (2002) also examines the applied mathematical subtleties of convex ‘*and*’ occurring in DJCCs, e.g. ‘*Fido and Spot are black and white*’. — Convex combination type connectives are available in developments of Fuzzy Logic, but the general framework of continuous truth- or membership-values is quite different from the present framework.

[22]From a calculus of n-ary properties $n \geq 0$ to a calculus of individuals supplemented by distribution meaning postulates as in the hybrid semantics.

[23]This much is already suggested by the familiar canonical mappings in multilinear algebra that yield correspondences of form $\mathbf{Hom}(A \otimes B, C) \simeq \mathbf{Hom}(A, \mathbf{Hom}(B, C))$, and so do familiar instances in categorial syntax and in logic as well as generalizations in category theory. Throughout, the linear spaces being assumed are over $\mathbb{R}$, $\mathbb{Q}$, or (for some smaller language fragments) $\mathbb{Z}$. As mention of syntax and logic in-

dicate, the abstract concept of multilinearity is much more general. For example (so Joachim Lambek found worth remarking in the 1990s), in the context of the theory of bimodules, Gentzen type sequents are referred to as multilinear mappings.

[24]The matter is indeed not simple. See below for an outline MLS treatment of negation, whose scope is well-known to be a downward monotonic context.

[25]Curiosity could not thus have its proverbial effect on this occasion.

[26]Subspace determination is quite analogous to that of a minimal requisite algebra for boolean semantics of a sentence or set of sentences.

[27]A reflex of the first is familiar from Otto Jespersen's early 20th century description of numeral usage. The explanation of the second, in which '*not*' can be glossed 'not only' is in Merin (1988, 1994). Concession is the 'marked' or non-default act in the dual pair. Categorical motivation for use of the term 'dual' is provided in Merin (1994). Marked contexts are characterized by reversed monotonicity. This holds the key to the phenomenon reported on p. 21 above: the re-tensed word string (1), whose assertoric default reading is synonymous with re-tensed (2), is reconstrued to receive a reading that is more analogous to re-tensed (3) in the embedded context exemplified on p. 21. That the predictive monotonicities are preference-induced, and not the consequence of a boolean semantics, will become plausible *prima facie* if one agrees that (i) substitution of '*applaud*' for '*get nervous*' increases one's inclination to assign a (2)-type reading to the embedded word string and that (ii) even just one of John and Mary singing or dancing will be understood to have neighbours getting nervous in the original example once the (2)-type requirement is relinquished. This reading of the string is less informative yet than the (3)-type reading predicted by KF or PTQ. Intuition (ii) is what warrants the description 'more analogous' as distinct from plain 'analogous'.

[28]Kracht (2007) is one subscriber to the claim. Gazdar (1979) coined the vivid slogan 'Pragmatics = meaning minus truth-conditions'. Pragmatics ought to include things to do with use and utility. As long as a pragmatics can be *a* semantics in the mathematician's liberal sense, I see nothing wrong with subsuming MLS under 'pragmatics'. All one should ask is that 'pragmatics' doesn't necessarily stand for what you do if logic, formal semantics, syntax, or phonology are too hard for you.

[29]Quite incidental to present purposes, there is a detailed argument in A. Merin & I. Nikolaeva (2008) 'Exclamative as a Universal Speech Act Category: A Case Study in Decision-Theoretic Semantics and Typological Implications', available online as <http://www.semanticsarchive.net/Archive/jUwMmYxY/eusac-sspa4.pdf>.

[30]To any expression for an MLS-vector denotatum in a real vector space there will be infinitely many denotationally equivalent expressions. Fact is that we and our languages know how to pick the relevant representatives of this equivalence class of expressions. The same (non-)problem also arises for semantics in the smallest of non-degenerate boolean algebras, never mind people-sized ones. The universe of symbols is full of junk, both sensible and non-sensical. The latter kind are elements of denotation spaces which are much like chimeras and are not denoted by anything anyone would care to express. We are untroubled by either kind when of sound mind.

[31]They are unlike Peirce's well-behaved prototype of what he called an 'icon', which was the sentence of a natural language or well-formed formula of a mathematical calculus.

[32]This is the law for which readers of Keenan & Faltz (1985:71) are invited to construct a validating example, after validation of its dual has been illustrated with a real sentence of English. Absorption and idempotence are likewise treated as empirically unproblematic.

[33]I use the 'either' in the NL examples as the scoping device it is. Indications of prosodic commas would do just as well.

[34]They would not, on this count, necessarily violate pertinent connective laws of resource-dependent substructural logics such as linear logic.

[35]Unlike '*A and A*', the '*A or A*' schema exhibits no mismatch in MLS between linear additivity and evidential idempotence. What is being violated is the equivalent of a use-condition by which late 19th century accounts of everyday '*or*' conservatively extended a boolean semantics for it. The speaker is required to be ostensibly ignorant which of the disjuncts is true or realized. If '*A or B*' is truthfully or felicitously asserted, at least one disjunct must be true or realized, which is incompatible with the ignorance condition when $B = A$. For imperatives, (e.g. '*Go to Rome or go to Rome!*') the addressee will, under the MLS construal, be given a formal choice that is, in fact, no choice. Iteratively readable verb phrases (e.g. '*walk and walk*') for which additive readings are plausible, lead to acceptable '*A and A*'-coordination and are dealt with in Merin (2002).

[36]Generalists will get an inkling of the relationship between games and linear algebras by recalling that von Neumann's (1928) theory of two-person constant sum games is essentially an application of the theory of bilinear forms.

[37]Vector space representations (for sets of query terms and documents) have found uses in information retrieval from the 1970s onwards. Regardless of applications domain, they arise naturally in statistical analysis of correlation.

[38]Another quite specific motivation is to let this be semantics for an algebraic theory of syntactic composition, Lambek's Pre-Group Grammars, which has grown out of the Syntactic Calculus a.k.a. Lambek Calculus (Lambek 1958). This calculus was also one of the earliest 'substructural' logics, of which linear logic is a more recent instance; and it was, in its turn, a spin-off from a calculus for multilinear mappings.

[39]Recall that there are as many potentially distinct possibilities as there are distinct value n-tuples for the selection functions of n occurrences of '*or*'. Here $n = 2$.

[40]Try other contents or the '*and-or*' dual form if you think, wrongly, that the infelicity is due to the disparate clause contents and attendant irrelevance of the utterance.

[41]A historic example of successful unification is the categorical approach.

[42] Sections 1 to 3 slightly expand a short paper 'Multilinear Semantics for Double-Jointed Coordinate Constructions' submitted for the IATL 22 conference in March 2006. As <http://semanticsarchive.net/Archive/jdjODVmY/mlstct.pdf> this text has been online with reformatted layout as of 2008. Research support by the Thyssen Foundation and the German National Science Foundation (Deutsche Forschungsgemeinschaft, DFG) is gratefully acknowledged.

References

Aristotle, *Sophistici Elenchi*. Ed. W.D. Ross, Oxford: Clarendon Press, 1958. [Trsl. W.A. Pickard-Cambridge, 'Sophistical Refutations', in J. Barnes (ed.) *The Complete Works of Aristotle* Vol. I, Princeton NJ: Princeton University Press, 1984.].

Asser, G. (1957). Theorie der logischen Auswahlfunktionen. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 3, 30–68.

Birkhoff, G. & von Neumann, J. (1936). The logic of quantum mechanics. *Annals of Mathematics* 37, 823–843.

Chomsky, N. (1955). Logical syntax and semantics: their linguistic relevance. *Language* 31, 36–45.

Gazdar, G. (1979). *Pragmatics: Implicature, Presupposition, and Logical Form*. London: Academic Press.

Girard, J.-Y.. (1987). Linear logic. *Theoretical Computer Science* 50, 1–102.

Grice, H.P. (1967). Logic and Conversation. Repr. with revisions in H.P.G., *Studies in the Way of Words*, Cambridge MA: Harvard University Press, 1989.

Hölder, O. (1901). Die Axiome der Quantität und die Lehre vom Masz. *Sitzungsberichte der Verhandlungen der Kgl. Sächs. Ges. der Wissenschaften, Leipzig, Math. Phys. Klasse* 53, 1–64.

Keenan, E. & Faltz, L (1985). *Boolean Semantics for Natural Language*. Dordrecht: Reidel.

Kracht, M. (2007). The emergence of syntactic structure.. *Linguistics and Philosophy* 30, 47–95.

Lambek, J. (1958). The mathematics of sentence structure. *American Mathematical Monthly* 65, 154–169.

Link, G. (1983). The logical analysis of plurals and mass- terms: a lattice-theoretical approach. In *Meaning, Use, and Interpretation of Language*, eds. R. Bäuerle, C. Schwarze & A. von Stechow, 302–322. Berlin: De Gruyter.

MacLane, S. & Birkhoff, G. (1967). *Algebra*. New York: Macmillan.

Melucci, M. & van Rijsbergen, C.J. (2011). Quantum mechanics and information retrieval. In M. Meluccci & R. Baeza Yates (eds.) *Advanced Topics in Information Retrieval*. Berlin: Springer, pp. 125–155+.

Merin, A. (1986). 'Or', 'and': non-boolean utility-functional connectives. [Abstract] *Journal of Symbolic Logic* 51, 850–851.

—— (1988). Sociomorph Quasi-Linear Semantics. Dissertation submitted for the Internal Junior Fellowship Competition, King's College, Cambridge.

—— (1994). Algebra of elementary social acts. In: S.L. Tsohatzidis (ed.) *Foundations of Speech Act Theory*. London: Routledge, 234-263+.

—— (1997). If all our arguments had to be conclusive, there would be few of them. *Arbeitsberichte des SFB 340* Nr. 101, Universities of Stuttgart and Tübingen. Online at http://www.ims.uni-stuttgart.de/projekte/SFB340.html and as <http://semanticsarchive.net/Archive/jVkZDI3M/101.pdf>.

—— (1999). Information, relevance and social decisionmaking: some principles and results of Decision-Theoretic Semantics. In L.S. Moss, J. Ginzburg & M. De Rijke (eds.) *Logic, Language, and Computation* Vol. 2. Stanford CA: CSLI Publications. pp. 179–221. Also available online e.g. as <http://www.let.uu.nl/esslli/Courses/merin/irsdmu.ps>

—— (2002). Linear Semantics: Values and Images. *Forschungsberichte der DFG-Forschergruppe 'Logik in der Philosophie'* No. 100, University of Konstanz.

Montague, R. (1970). Universal Grammar. *Theoria* 36, 373–398. Repr. in Montague (1974:222–246).

—— (1973). The proper treatment of quantification in ordinary English, In J. Hintikka, J.M.E. Moravcsik & P. Suppes (eds.) *Approaches to Natural Language: Proceedings of the 1970 Stanford Workshop on Grammar and Semantics*. Dordrecht: Reidel, pp. 221–242. Repr. in Montague (1974:247–270).

—— (1974). *Formal Philosophy: Selected Papers of Richard Montague*. Ed. and intr. R.H. Thomason. New Haven CT: Yale University Press.

von Neumann J. (1928). Zur Theorie der Gesellschaftsspiele. *Mathematische Annalen* 100, 295–320.

Preller, A. & Sadrzadeh, M. (2011). Semantic vector models and functional models for Pregroup Grammars. *Journal of Logic, Language, and Information* 20, 419–443.

van Rijsbergen, C.J. (2004). *The Geometry of Information Retrieval*. Cambridge: Cambridge University Press.

Widdows, D. (2004). *Geometry and Meaning*. Stanford CA: CSLI Publications.

Winter, Y. (1998/2001). *Flexibility Principles in Boolean Semantics: the Interpretation of Coordination, Plurality and Scope in Natural Language*. Ph.D. Thesis, Utrecht 1998; Publ. Cambridge MA: MIT Press, 2001.

Zwarts, F. & Winter, Y. (2000). Vector space semantics: a model-theoretic analysis of locative prepositions. *Journal of Logic, Language, and Information* 9, 169–211.

Author address:
Department of Philosophy
University of Konstanz
78457 Konstanz
Germany
arthur.merin@uni-konstanz.de