



Exemplar theory in linguistics: a perspective for the cognitive subject

René-Joseph Lavie

► To cite this version:

René-Joseph Lavie. Exemplar theory in linguistics: a perspective for the cognitive subject. 2007.
halshs-00142394

HAL Id: halshs-00142394

<https://shs.hal.science/halshs-00142394>

Preprint submitted on 18 Apr 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Exemplar theory in linguistics: a perspective for the cognitive subject ¹

René Joseph Lavie

University of Paris 10-Nanterre, France

1 Introduction

In linguistics today, some amount of attention is devoted to exemplar-based theories and exemplar-based dynamics¹. This approach was first taken in phonology, where it is still only used by a minority but is gaining in popularity. The theory of exemplars is initially a theory of perception in psychology. It has been applied to the perception of the sounds of discourse, notably by Johnson (1996). It constitutes a good model of the categorisation of vowels. The model of Hintzman (1986), which is close to it, has been applied by Goldinger (1996) to word recognition.

Exemplar-based models share several features with neuromimetic connectionist models without however coinciding with them². They shift attention away from abstractions towards individual, occurrence-based facts; they reduce or nullify the importance of categories, rules and constraints³ and on the contrary they stress the dynamics, that is, the processes. In spite of an impoverished explanatory apparatus, with respect to linguistic theories which are symbolic in their majority, exemplar-based models succeed in accounting for categorisation effects and regularisation effects and they achieve numerous explanatory successes.

I introduce briefly the theory of exemplars in phonology following a particular model, that of Janet Pierrehumbert. Then I show how the exemplarist track may be accommodated into syntax by referring to my own work. In the course of the paper, I stress several aspects that show why such approaches have cognitive relevance.

In conclusion, I suggest that an exemplar-based theory of the cognitive subject could benefit from notions elaborated in linguistics⁴.

2 Exemplarism in phonology

Pierrehumbert (2001) summarises the theory of exemplars as follows. A category (for example a phoneme) is represented by a cloud of numerous exemplars (or tokens or instances) of the category. Strongly similar instances are close together. Stored exemplars present all the variation that occurs in the physical correlates of the category. For example, the exemplars of the vowel /ε/ present the variety of formant values (anatomy of the vocal tract, hypo-hyperarticulation, etc.): variation of duration, of F0⁵, etc. The phonetic parameter space is mapped onto the labels⁶ of the category system. The number of exemplars per category helps explain frequency effects.

Viewing the memory as constituted of exemplar clouds raises the objection that the memory capacity of an individual is not sufficient to store all episodes *in extenso*. Pierrehumbert envisages the following possible responses: a) memory decay (recent memories against more ancient ones), b) the parameter space (duration, F0 for example) is granularised, that is, the coding does not differentiate too small differences.

Each stored exemplar has a strength which is a persistent activation. The more recent and more frequent exemplars are stronger. In figure 1, each vertical line represents an exemplar; the dashed lines are the exemplars of phoneme /ε/ and the continuous lines are the exemplars of phoneme /i/. The height of

¹ Communication to the 11th Congress of Cognitive Linguistics, Bordeaux, 19-21 May 2005. Proceedings in press.

each bar is proportional to the strength (or activation) of the exemplar. For each bar, the abscissa is the value of a physical parameter associated with the exemplar: F2. This is an introductory simplification: an exemplar actually has several parameters attached to it; the parameter space is pluridimensional.

When a new element is met and has to be categorised (in figure 1, the new element is the star which is at the center of the double arrow): a) the new element undergoes a perceptual encoding which assigns it a place in the parametric space, b) some exemplars are identified as the closest according to their distance in the parametric space, c) then the closest exemplars are considered and, depending on their labels, the most probable labels for the new element are worked out. The F2 of /i/ is generally greater than that of /ε/, but the exemplars of both phonemes have an overlap (dialect differences, speaker idiosyncrasy, production biases such as coarticulation). So the vowel to categorise is ambiguous between /ε/ and /i/ but phoneme /i/ wins because it has more exemplars being activated in the neighbourhood of the star.

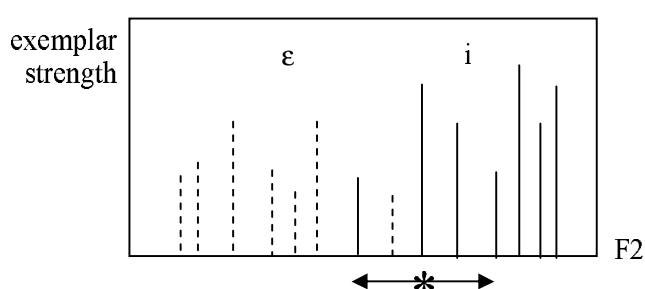


Figure 1 A new element has two possible labels : /ε/ or /i/

More frequent labels have more exemplars being activated so they have more influence in ambiguous cases. Thus the model predicts a bias in favour of frequency, which is consistent with empiry.

The temporal dimension is absent (that of the realisation of the acts). Decision rules in this model are only a substitute for an activation/inhibition system, which is rather assumed in the reality to be responsible for activating the label's activation by activation propagation: the winner then would override the activations of its competitors. However, this model is coherent with the idea that phonological decision time, and lexical decision time, reflect the time that it takes to reach decision thresholds.

The advantages of this approach are the following ones (I still summarise Pierrehumbert):

- The acquisition of phonetic and phonologic knowledge is the acquisition of a large number of exemplars. The model has a better descriptive adequacy than its competitors. A universal symbolic alphabet (Chomsky & Halle 1968, Chomsky & Lasnik 1995) does not represent the fine differences which languages make between the values and the distributions of phonetic properties and it does not tell how they are acquired.
- A token located at the center of a cloud is recognised in the category, even if there is no exemplar for the category exactly at the center of the cloud. Thus there is no need to reify a prototype.
- For a phonological category, extreme exemplars are sometimes judged better than ones located at the centers of the cloud. Thus for vowels /i/ and /u/ the best exemplars, according to speaker judgments, have more extreme formant values than the values associated with more typical productions. Here is why the model accounts for this observation: seeing that its attribution of a category is contrastive with others, the distance of a candidate from all the exemplars of competitor categories increases its subjective goodness. What counts is not that it is a good /i/, but rather that it would be a very bad /u/.

- One accounts for frequency effects with an inherent property, rather than having to make frequency an implausible add-on to a base-theory that would ignore frequency. Exemplar-based theories can be complemented by assumptions about the neural implementation and thus explain how frequency influences their productions and their speed.

Building on this basis, Pierrehumbert then develops four models of phonological production (thus far, exemplar-based modelling has rather been applied to reception):

A first model (model A) accounts for the ability to produce variant realisations of categories (variation of occurrence-based productions and inter-speaker variation) and to recognise them despite variation. Therefore, it accounts for categorical robustness between speakers despite the variation of category realisations. This model presents no systematic bias. Now a systematic bias is necessary to account for the tendency to under-articulate by economy.

A second model (model B) is then derived from the first one by introducing a bias. It accounts for historical lenition⁷ by linking it correctly to frequency (in English, in *every*, *evening*, which are frequent, the second *e* is hardly to be heard, whereas in *mammary*, *artillery*, which are rare, the second *a* of *mammary* and the *e* of *artillery* are explicitly realised). The model, modified in this way, predicts the following (and is the only one to predict it): i) each word has a production variability; ii) the effect of frequency on lenition rate is gradient; iii) the effect of frequency on lenition rate must be observed in individual discourse (a community of speakers is not an average between the generations within it); iv) the synchronic observation of the pronunciation of words of various frequencies and the diachronic observation that the pronunciation of a word by a person changes over time. The model also predicts that the distribution must have a bias, small or significant depending on the change rate. This could be checked only with numerous data which we do not have. A speaker moving to a different linguistic environment changes his pronunciation schemas only slowly (months or years). Historical changes take decades or centuries. Thus the high iteration numbers in the model ($\sim 10^4$) are realistic. Historical changes affect the old less than the young and so a variance is attached to the coexistence of generations. To this, the model suggests two possible factors: i) for a given pattern, the old have more exemplars than the young and so older exemplars dominate (this explanation is valid only if the model of oblivion is a slow decay); ii) the old add fewer new exemplars because their memory lends itself less to it; therefore their clouds are more dominated by ancient exemplars.

The systematic bias introduced in model B however has the consequence that exemplar clouds spread with time, and this increases the variance of the realisations; such increase is contradicted by empiry: on the contrary, entrenchment decreases the variance. A third model (model C) corrects this by modifying the way productions are assigned random variation: random variation no longer consists of choosing an exemplar at random but it now consists of choosing a point in the parametric space at random. And integration is a result of this change because the new approach tends to move closer to the distributional average.

Finally, a last extension of the model (model D) succeeds in accounting for the neutralisation of a phonological opposition⁸: a marked category, when it is close enough to an unmarked one, ends up with time to be assimilated by it.

Pierrehumbert concludes that the exemplarist dynamic, which supposes the memory constituted of labelled exemplars, reproduces several results of usage-based phonology:

- the possibility to learn the phonetic patterns of a language,
- the slow evolutions of adult speech,
- the fact that lenition is more marked in more frequent words,
- the phenomenon of neutralisation, of which the model provides a precise model.

This approach has a fundamental interest. It supersedes prototype theories⁹ because it integrates inherently and centrally production variability, inter-speaker variation, and diachronic evolution.

In my view, the question of oblivion in it calls for a comment. The assumption of memory decay at the expense of older exemplars¹⁰ is not sufficient because: a) of pathology; senescence and the Alzheimer syndrome show on the contrary that recent memory is more fragile in certain conditions, b) of psychoanalysis which teaches us that certain configurations have no plasticity (which would constitute the psychoanalytical unconscious) and c) of the early acquisition of mother tongue phonology and of the resistance thereafter to acquire a different one. Pierrehumbert herself sets forward a different proposition at the end of the paper: older speakers would learn more slowly and their knowledge (for me, their know-how) would be dominated by more ancient inscriptions.

The model encompasses a critique of categories in two respects: i) their characterization by properties, or the 'descriptive' approach to them to adopt Millikan's word¹¹, is abandoned, and the evolution of their system is explained (the system is not postulated fixed, out of time). Still, the model rests critically on categorical labels (the phonemes). We ought to be cautious, generally about categories, and specifically in a domain, phonology, in which the focus shifts back and forth from syllable to phoneme every two decades or so. We should notice however that the reference to categories, which phonemes are, does not constitute a direct flaw of the theory of exemplars because the model proposed by Pierrehumbert touches here the limit of its perimeter: its extension is just up to the phonemes. If it were prolonged so as to encompass syllables and words maybe it could do without categorical labels (phonemes at any rate). What then could be a possible extension: a) to 'words' assembling several phonemes, b) to longer syntactic combinations?

3 Exemplarism in syntax: the 'Analogical Speaker'

To my knowledge, two approaches apply to syntax which may be said to be exemplarist. The first one is that of Skousen (*et al.* 2002) which is founded on massive computations applied to corpora; in my opinion, its cognitive relevance is questionable (but the approach is interesting in other respects) and it is not appropriate to go into more details about it in this article. The second one is the track I followed in my doctorate work.

In order to model the linguistic fact (I do not write 'a language') I showed¹² the shortages of the grammatical approach: a set of stipulating and static¹³ propositions (a grammar) accounts for a non-variant and fixed 'language' always with residues. Further, it meets difficulties in accounting for inter-speaker variation, acquisition, and thence linguistic change.

In order to cope with the theoretical difficulties met by symbolic and categorical theories facing variation, acquisition and linguistic change, a stream of thought¹⁴ has developed in the nineties aiming at adding to such theories a stochastic complement. These propositions do obtain some descriptive enhancements but they have not succeeded in causally linking in a convincing way either of the three to the acts of utterance reception, and utterance production. So they stay with a worrisome explanatory breach¹⁵.

In short, making the assumption of a 'language', even understood as an 'internal language' or an 'individual language' is not the best track to take because a speaker does not learn a language, a speaker rather learns how to speak. If we accept this, the model or theory of what is in a speaker's head which enables him to succeed in his linguistic environment must, right from the start, conjoin a model of dynamics¹⁶ with a static model. A static-only model or theory is not sufficient. A 'language' as an entity which can be described and studied for itself is an object for history or sociology but it is not something which resides in our heads and the elucidation of which would be a prerequisite to the explanation of the linguistic dynamics. By addressing the static and the dynamics separately, the approach used to be disjunctive; it must become conjunctive.

4 Architecture of the Analogical Speaker

We need however hypotheses: i) about the permanent mental traces in a speaker's head which support the dynamics, and: ii) about the dynamics themselves. Such assumptions cannot be verified or falsified directly; they can only be observed indirectly by their consequences in a precise model which we can build on.

The mental traces are organised into a 'plexus', so named because it is strongly meshed, and the dynamics are abductive and implemented by a network of agents.

Inscriptions in a plexus are analogy-based but the analogy at stake is understood as being a dynamical one. The recognition of analogy as static ratios obtaining between terms and as dynamic processes active in linguistic acts restores a continuity – well under looked in the 20th century – with two thousand years of linguistic thought, from Varro until Householder (1971). The Neogrammarians and Saussure identified indeed a dynamic analogy which 'repairs' the paradigms damaged by phonetic change; they caught sight of the operative schemas which would support this without however developing them. Several linguists, among which is Bloomfield, also assumed that analogy has the power, in synchrony, to explain individual speaker acts, and acquisition, but here again without proposing explicit mechanisms. The value of analogy was posited, but one would not go as far as motivating explanatorily what speakers do and do not do¹⁷.

The model proposed here acknowledges Neogrammarian and Saussurean theses about the function of analogy in diachrony, it takes over Bloomfield's intuition about its function in the acts of emission and reception, and it undertakes – what they did not do – the building of a precise and operable base to their intuitions¹⁸. The model in question was not possible at their time: its underlying theoretical propositions can be falsified only indirectly through their consequences and this encompasses a great number of terms and elementary operations so that validation must involve a computer implementation. In this view, is falsifiable a proposition which is precise enough to enable the construction of a model and it is validated – with other ones – if the model behaves in some way as natural speakers would.

As an example, here is a set of analogical inscriptions¹⁹; they are structure mappings.

<i>willing</i>	<i>to</i>	<i>please</i>	<i>willing to please</i>
<i>eager</i>	<i>to</i>	<i>win</i>	<i>eager to win</i>
<i>willing</i>	<i>to</i>	<i>walk</i>	<i>willing to walk</i>
<i>trying</i>	<i>to</i>	<i>understand</i>	<i>trying to understand</i>

The four assemblies of the rightmost column are mapped onto one another according to the alignments shown by the three first columns. Such a set of inscriptions²⁰ reflects the assumption that speakers recognise these constructions as similar (formal similarity and semantic similarity, I come back on this below).

Between inscriptions the model encompasses a proximity which is adherent to a notion of access. Starting from a defined inscription, a few more may be accessed at low cost, then, in further steps, more can be reached but at a higher cost. In the example, the exemplar *willing to please* (line 1) and the exemplar *eager to win* (line 2) are directly linked – in the jargon of the model, there is a 'paradigmatic link' between them – one is accessible from the other in one computation step. Similarly between line 2 and line 3, same again between line 3 and line 4. But there is no direct access from 1 to 3. Access can happen however from 1 to 3 in two steps, via line 2: if 3 is analogous to 2 and 2 is analogous to 1, then, abductively, 3 is analogous to 1. This constitutes an 'elementary movement', which is abductive, it is the 'movement by transitivity', one of the four movements (below) whereby the static inscriptions are exploited by the dynamics.

Proximity is not so much associated with the inscriptions themselves – the statics on its own defines proximity in part only – but rather with the use which the dynamics (below) make of the inscriptions.

Stress on proximity fosters the view that what matters is not the inscriptions in their hypothetical 'essence' but rather accessing them, then re-accessing more inscriptions from ones already accessed; that is, the focus is displaced on transitions.

If one may globally accept the structure mappings above, one may also find that there is no definite reason to connect lines 1 and 2 together rather than lines 1 and 4 for example. Without further condition an important arbitrariness seems thus to be associated with plexus structure. It reduces when testing the plexus and tuning it to yield a behaviour acceptable for a speaker of English; but the arbitrariness does not reduce down to zero. The proposition is to regard the remaining plexus arbitrariness as accounting for speaker idiosyncrasy (section 8 below).

Turning now to the dynamics, a first assumption is that they are abductive in nature and rest on four basic 'abductive movements' which lead the static data into the dynamics. We just saw on the example the first movement, called 'by transitivity'.

A second abductive movement is by 'constructibility transfer': in the inscriptions given above, the model avails itself to regard as abductively possible for this speaker to utter or receive assemblies like *willing to win*, *trying to walk*, then *trying to please*, *willing to understand*: it commutes terms, first between adjacent lines, then between more remote ones.

A third abductive movement is 'by expansive homology'. It may take place when a configuration of the plexus makes homologous together: a) a term, and b) an expansion of the term²¹. Here is an example, again an excerpt of the plexus which will support the case in section 4.

<i>do</i>	<i>talk</i>	<i>do talk</i>
<i>don't</i>	<i>talk</i>	<i>don't talk</i>
<i>don't</i>	<i>talk to him</i>	<i>don't talk to him</i>

Term *talk* and term *talk to him* are homologous, and *talk to him* is an expansion of *talk*. This – along with a few more inscriptions – makes it possible for the model to abductively license forms like *do come back to me*, *don't come back to me*, etc. This is the lever of what is sometimes called 'language recursivity'.

A fourth abductive movement concerns the pluridimensional paradigms which are frequent in linguistics. Thus, beside structure mappings – the 'structural analogies' above – the model encompasses a second type of analogy: 'systemic analogy' as for example²²:

- (1) *villes : ville :: campagnes : campagne*
- (2) *mangerons : mangerai :: mangerez : manges*
- (3) *la : le :: une : un*
- (4) *soigneux : avec soin :: rapide : vite.*

Systemic analogy may be sanctioned by linguistic form (line 1), be sanctioned partially only (lines 2 and 3), or not be sanctioned at all (line 4). Systemic analogy is important for agreement, concord, and more questions of non-adjacent dependency. Its integrative application makes it possible to explain how, in the infant speaker, the productive reinstallation of the same paradigms (e.g. the verbal paradigms of French) as those which he inherits from his parents takes place. About the same paradigms are reinstalled, not necessarily exactly the same ones. This approach, including such marginal variation: a) frees us from the *parameter setting* for example of the Principles and Parameters theory of Chomsky, and b) allows us to explain why speakers of French in the 21st century no longer use what was categorised as 'supine' or 'neuter' in 1st century Latin, there having been no sharp discontinuity of usage in the meantime notwithstanding²³.

Abductive movements specify elementary computation steps the combination of which makes it possible to reconstruct the linguistic acts. They are cognitively founded (or foundable), their implementation by neurons is more plausible than it is for rules, categories, or optimalist constraints.

Abductive movements are set at work within a general mechanism of abductive solving which deploys the heuristic tracks made possible by proximity, spreads them gradually and reacts to 'settlement conditions' which are detections of matches. The whole is inspired from tracks explored in AI – the part of AI which distanced itself from symbolism and cognitivism, somewhat in Minsky's way – and of the results of psycholinguistics. The exploitation of proximity, which is fundamental in the model, allows us to associate each act with a computational cost which is held to correspond to the cognitive cost: the analysis of an utterance with a given plexus has a given cost; and a same utterance will be received by different plexuses with different costs; note that the utterance's length is not the sole cost factor: paradigmatic distances, and the number of computation steps needed to reach the licensing exemplar(s) also matter.

5 Three corollaries of analogy

The Analogical Speaker contains three important principles which are corollaries of the option to ground the model on analogy and of way to do it.

The first one is the refusal of primarism. The rationalist tradition in science in general leads us to break our objects of attention into elements from which, by recombination and interaction, the immense variety of empiry is reconstructed. Success was great and unarguable: the elements of chemistry, then the neutron, proton, etc. However, recent tracks in fundamental physics, including the string theory, no longer follow these lines. In linguistics, rationalist primarism gave rise to the notions phoneme, lexeme, seme – I collected up to twenty-two words in *-eme* in authors of the 20th century. Here again, success was great, the phoneme, to say nothing of the rest, represented a substantial progress. However, these notions do not resist a finer scrutiny of the phenomena: accidents of compositionality limit confidence in the seme, morphophonological phenomena or ones of phonosyntax relativise the explanatory value or simply the descriptive value of the syllable or of the phoneme. Distributional analysis itself is not always possible and in a number of cases one simply fails to produce defendable morphemes. In all the compartments of linguistics we thus find effects which more suggest contingency (interaction of multiple factors, easy to conjecture and difficult to know, matters of fact, the contingent products of a history) than nomological necessity. This is why the orientation is taken to refuse by principle any foundational level where the 'quarks' of linguistics would form an apparatus, the recombination of which could account for everything. Since dynamic analogy seems to allow us to reconstruct productivities in the absence of such foundational postulations, they are simply not made.

The second of the three principles is that of the vacuity of terms. In the Analogical Speaker, a 'term' is whatever plays a role in an analogy. In the structural analogies (the structure mappings) above *willing*, *to*, *willing to please*, etc. are terms. As for a systemic analogy like *la : le :: une : un*, it holds between terms *la*, *le*, *une* and *un*. The same terms may occur in structural analogies and in systemic analogies; this ubiquity is an important factor of the model's integrative productivity. Terms thus defined must be free from properties like: a lexical category, an allocated lexical meaning, a reified argument structure, syntactic features, etc. Terms are 'punctual' so to say, without content, and their sole function is to manifest the recurrence of some identity. The efficiency levers which other theories obtain through properties are obtained here by the mutual placements of the terms in analogical inscriptions, that is, by their 'copositionings' (which are not relations). It is so because the attribution of properties to 'language individuals' is one of the reasons for the categorical frictions which compromise the required flexibility and gradience of effects. Moreover, granting properties to terms then makes the discourse about learnability very difficult: excessive demand made to innateness, causal chains with the acts badly built. Properties – if there had to be properties at all – should only be effects of the dynamics; they are not postulated in the apparatus itself. Thus terms are punctual, property-less and they take value and effect only from their recurrences in the analogies inscribed in the plexus, and from the copositionings they so acquire among themselves.

The third and last principle concerns contextuality. I remind (Greimas, Rastier) that polysemy is, for most of it, a side-effect of decontextualisation. A theory which begins with decontextualising, for example by postulating lexemes, then must face, and solve, difficulties which it itself just created and which do not belong to its object: the *factum linguae*. Constitutional contextuality (the refusal to decontextualise, contextuality 'right from the start') therefore appears as a desirable quality for a linguistic theory; it is however not easy to obtain, and it is very little obtained in the frames so far proposed. Here, the option of making inscriptions in a plexus only based on analogies, as explained above, ensures a certain contextuality. In a structural analogy this is manifest: *été* is never inscribed on its own, it occurs only in inscriptions like: *l'été prochain* (next summer) or *a été entendu* (has been heard). It is also possible to see a contextuality constraint in systemic analogies, though at the expense of an extension of the notion of context, we shall not go into details here. The principles of the Analogical Speaker ensure then a certain obligatory contextuality. It may be found partial and insufficient, no doubt it will have to be enhanced but, as it is, it allows us to reconstruct a substantial part of syntactic productivity while eschewing descriptive dilemmas such as decisions about categorical labelling. Hopefully, an extension of the model will be possible which will also allow us to eschew these other dilemmas that are those of the lexicographers (meaning 1, meaning 2, or else a generic meaning) and those of numerous semanticists (proper meaning vs. figurative meaning). Making contextuality obligatory in this way sets polysemy to its exact place: in context, it is very rare finally.

6 John is too stubborn to talk / to talk to / to talk to Bill

To provide a concrete feeling of how this model is efficient – but in the format of this paper I can only sample – I will now report briefly the treatment of a case raised by Chomsky (1981/1984: 13), which is a classical example in generative grammar. It consists of the three utterances below in which the agent of *talk*, changes two times as the utterance is prolonged with more words.

	Utterance	Agent of <i>talk</i>
(1)	<i>John is too stubborn to talk</i>	John
(2)	<i>John is too stubborn to talk to</i>	somebody else than John
(3)	<i>John is too stubborn to talk to Bill</i>	John

In order to do justice to this concentration of complexity, a model cannot commute in approximate conditions, it has to be very precise in the account it takes of something which underlies these utterances. In generativist propositions, this is the phrase structure. It is postulated, explicit, and its very definition supposes grammatical categories for terminal points (lexemes, morphemes) and categorical labels for syntagms²⁴. Without these, its definition could not even be stated. There is no intent here to deny the phrase structure: something of that kind is obviously at work in the dynamics of language acts. I rather undertake, abstaining from reifying it, to render its effects with simpler theoretical postulations.

The track adopted here is based on the recognition that speakers make easily a structure mapping between (1) and (3) but not between (1) and (2). This leads to accept in a plexus two different paradigms – in the sense defined above, i.e. two different chains of structure mappings – which remain unconnected or little connected, like P02+ and P02- below:

P02+ [the agent of V2 is the subject NP of a V1 expected on the left]

too stubborn	to	talk	too stubborn to talk
too lazy	to	work	too lazy to work

P02- [the agent of V2 is not the subject NP of a V1 expected on the left]

too big	to	take away	too big to take away
---------	----	-----------	----------------------

too difficult	to	understand	too difficult to understand
too difficult	to	please	too difficult to please

Such inscriptions do not mix up cases in which the agent of the second verb is that of the first one with cases in which it is not. It is then possible²⁵ to set up a productive model in which these separations of agentive orientation are propagated with accurateness to novel linguistic facts. The utterances (1), (2) and (3) are analysed successfully. In another example, the utterance *John is eager to please* is licensed by a structure mapping with the utterance *Alice is willing to work*, and not with *French is easy to learn*. The 'licenser' (*Alice is willing to work*) has the same agentive orientation as the analysed utterance. This property obtains with many examples; the model is robust.

The apparatus necessary to yield such results is that which was briefly sketched above: plexus inscriptions (like paradigms P02+ and P02-), abductive movements by transitivity, by constructibility transfer and by expansive homology; all this triggered in an abductive solving mechanism. Each of these items makes modest theoretical demands, and stays close to a certain implementational plausibility. It does not depend on a particular language or on a particular kind of linguistic structure. Currently, it has some limits though about which I shall provide some details below.

It is appropriate to stress again that these results obtain without any abstraction like category, rule or constraint: the model is fully exemplarist. If it still contains an abstraction at all, it is the dynamic itself of abductive solving. This would be a universal – but it is not a grammar.

7 Statement of achievements and an assessment

The example just studied is typical of what the Analogical Speaker can yield in its current development status. It is capable of an important part of syntax and of inflexional morphology. It has currently two main limits.

In the first place, the model does not cover grammatical agreement nor some more non-adjacent dependencies; the model is too narrowly distributionalist, too 'Harrisian' so to say. It would be fairly simple to make it agreement-capable by introducing syntactic features as do chomskyan and post-chomskyan approaches. However this would constitute an addition to the theory which would be categorical and consequently too heterogeneous to the rest of its postulations. Moreover the thesis on learnability (cf. below) would then become confused and less strong. This course is not taken and research on non-adjacent dependencies is pursued in different directions.

The rejection of syntactic features must be connected with the question of 'systemic productivity'. In this approach, linguistic productivity is reenvisioned; there are two modes of productivity actually: structural productivity²⁶, which is well identified, well described, and explained to some extent by this theory and other theories, and systemic productivity²⁷ which, in my view, has been wrongly approached, cognitively speaking – when it has been approached at all, a great deal of theories just do not touch it. The solution to it that calls on syntactic features is firstly descriptively inadequate because it accounts sufficiently neither for the cline anomaly-regularity of the forms, nor for anomalies in the paradigmatic frames themselves. But most importantly it is inadequate because it is cognitively impossible to found: feature systems are not learnable in a manner which makes room for inter-speaker variation and for language evolution²⁸.

The second important limit of the Analogical Speaker in its current development state is that meaning matters are not really addressed in it. The format of this article does not even allow to sketch the research tracks in this direction; I just mention that neither lexical meaning nor linguistic meaning seem to have to be posited as given and causal in an interpretation, and that a constructivist and interpretive approach is favoured.

Within these limits however, numerous properties externally perceived as 'grammatical' can already be reconstructed as results. The general argument is that grammatical stipulations are only constructed consequences of the dynamics, counter to e.g. the generativist claim that a static theory of a language: a) is doable, and b) is previous to the explanation of the dynamics.

Two aspects are now going to be detailed because they have a particular cognitive relevance: the reconciliation between speaker idiosyncrasy and linguistic quasi-normativity (section 8) and a model of acquisition which is possible in this frame (section 9).

8 Determinism / idiosyncrasy / quasi-normativity

The structure of a plexus, that is, the precise detail of the inscriptions in it, poses to the reader – and before him also to the descriptor – the question of the residual arbitrariness attached to it; we already met this point above. 'Arbitrariness' here is not the arbitrariness of the sign (its conventionality), but rather questions like "Why inscribe this term and not this other one; for a given term, why in this record and not in another, why these particular paradigmatic links". Such questions may have occurred to the reader on the occasion of the paradigms provided as examples above. This arbitrariness is not as residual as that: when some obvious description needs have been satisfied (such term is a must, such inscriptions are obviously closer together than such other ones, etc.) a great deal of description microdecisions still remain, and have to be made with no particular reason. The descriptor then makes an arbitrary choice. Ensuing tests with the dynamics generally suggest corrections which are a way to move to a new, better motivated status of the plexus. However, even after validation and correction, the motivation is far to command the entire plexus detail and a great deal of arbitrariness remains. It is important not to leave it without an interpretation.

A part of the plexus arbitrariness may be put on behalf of the radical exemplarist assumption made in this research: as it is too radically poor in its apparatus, the model is too unspecified; a less radical model, but which remains to be found, would be tighter and its inscriptional detail more constrained.

I propose to see the rest of plexus arbitrariness as standing for the speaker's idiosyncrasy: a plexus, being the static side of a speaker's linguistic knowledge, bears the trace of the speaker's history, of his learning history in particular. The licensing of a new utterance, as we saw above in the example *John is too stubborn to talk*, is achieved by means of an exemplar inscribed in the plexus (or of some exemplars more generally). In two different plexuses, representing two different speakers of 'the same language', the precise, exemplarist, conditions of the licensing are different, because the linguistic and cognitive histories are different and so are the permanent traces they left. Despite that, about the same utterances are licensed with about the same cognitive costs – in the model, the analogue of the cognitive cost is the number of computation steps required – and the two (models of) speakers 'speak the same language' with the variation which is to be observed in nature. To achieve this, the critical theoretical point is not to postulate a language; not to postulate a language like French of course, but, even more importantly, not to postulate an individual language.

Figure 2 proposes a metaphorical view of the question. It suggests that an important detail variation is damped and gives quasi-uniform linguistic outputs (they are quasi-normative), that is, the linguistic knowledge of this speaker is French, for example, with the variation across speakers which is to be observed among French speakers. The damping would be accounted for in general by a form of stability in complex systems and in particular by the intergrativity properties of the model (cf. Lavie 2003:229).

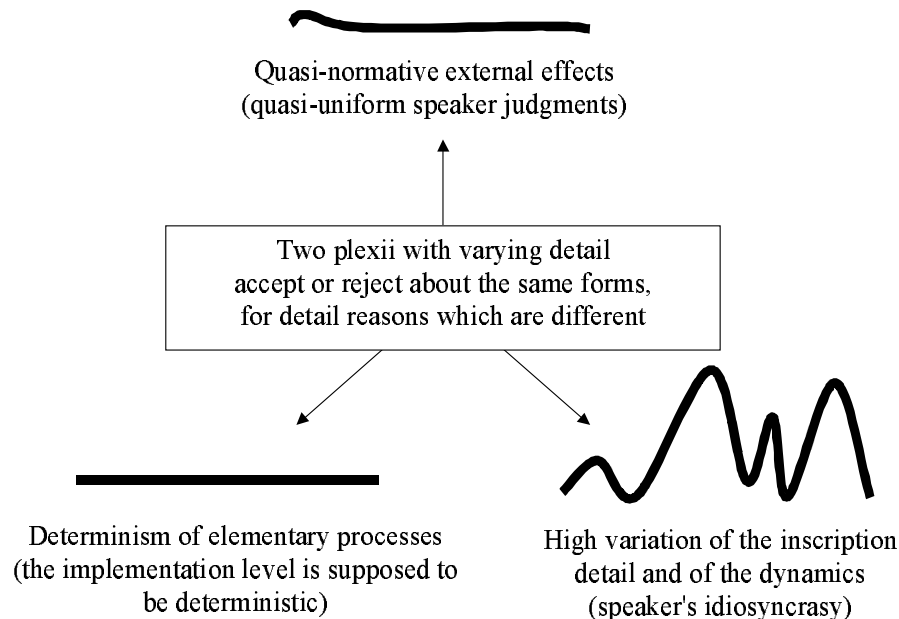


Figure 2 Determinism, idiosyncrasy, normativity

This schema reconciles three poles:

- a) the high variation of inscriptional detail across speakers (and of the detail of the dynamics) which is assumed to reflect the idiosyncrasy of speakers and the occurrence-based differences in individual histories,
- b) the quasi-uniformity of macroscopic effects²⁹ and
- c) the determinism of the neurophysiological processes which is assumed. The neurophysiological processes which support the linguistic operation have to be deterministic if we think that they belong to chemistry and therefore do not require calling on quantum mechanics. This is a conjecture but in any case, noise in synapses for example cannot explain interindividual variation in tension with the quasi-normativity of the productions.

The three-pole model above contradicts the explanation of several variation effects by means of probabilities³⁰. It also affects the theme of portability and separation. For Putnam (1960), the fact that a same process may be run on different computers (or, more abstractly, that a Turing machine is a "logical" description which leaves undetermined its concrete form) leads to envisage mental states and mental processes that can be described separately of the nervous system. This important remark is presented as likely to solve the problem of body and mind³¹. It legitimates a theme which is central in cognitive science: the postulation of a representation level independent of the hardware.

The three-pole model proposes a less sharp vision of this. First it does not posit an abstract object (it would be a language) which would be portable: the speaker productions are quasi-normative, they are not normative. Secondly, idiosyncrasy (bottom right pole) is both the variant result of an individual history and a dependency on the "hardware". The separation then could take place only at the expense of an abstraction (the postulation of a language) which we are trying to avoid. If one posits the possibility of a

separation, one cannot provide an explanation of variation or a working explanation of learning because it cuts the model off from the concrete dynamics of the acts.

9 Acquisition

The Analogical Speaker encompasses a simple acquisition model. It amounts to positing that a new utterance, upon its licensing by an already inscribed exemplar a) becomes inscribed itself in the plexus as a new exemplar, and b) connects to the licensing exemplar by a 'paradigmatic link' reflecting the structure mapping between them. This model is compatible with an acquisition model once proposed by Minsky (1985:82), the theory of 'Knowledge lines' or 'K-lines': *the assumption that we keep each thing we learn close to the agents that learn it in the first place.*

It explains the progressive generalisation of a usage in the learning infant. The constant empiry of psycholinguistic experiments is that, in an infant, a syntactic novelty never appears as suddenly available in the entirety of its potential application domain. Its spreading rather follows an epidemic propagation schema as illustrated below, where italics represent the new device – here the arrival of the copula – the usage of which becomes progressively more general:

Initial state	I n t e r m e d i a r y s t a t e s		Final state
Daddy gone	Daddy gone	Daddy gone	<i>Daddy is gone</i>
Jo naughty	Jo naughty	<i>Jo is naughty</i>	<i>Jo is naughty</i>
cat dead	<i>cat is dead</i>	<i>cat is dead</i>	<i>cat is dead</i>

Figure 3 displays a result taken at random in the literature (van Kampen and Evers, 2000). Here the phenomenon is characterized in the formalism of the Government and Binding theory but this is not the point that matters.

A syntactic acquisition by a speaker first appears as one or a few occurrences after what its extension grows progressively, slowly at first, then faster, then slowly again gaining finally the last few exemplars. The process follows a sigmoid curve and its time span lasts 20 to 40 weeks depending on the phenomenon and on the speaker. For a given speaker, several such sigmoidal acquisitions will succeed in time, massively between two and three years, but each instance of the generic process lasts between 20 to 40 weeks. All reported empirical results follow this schema.

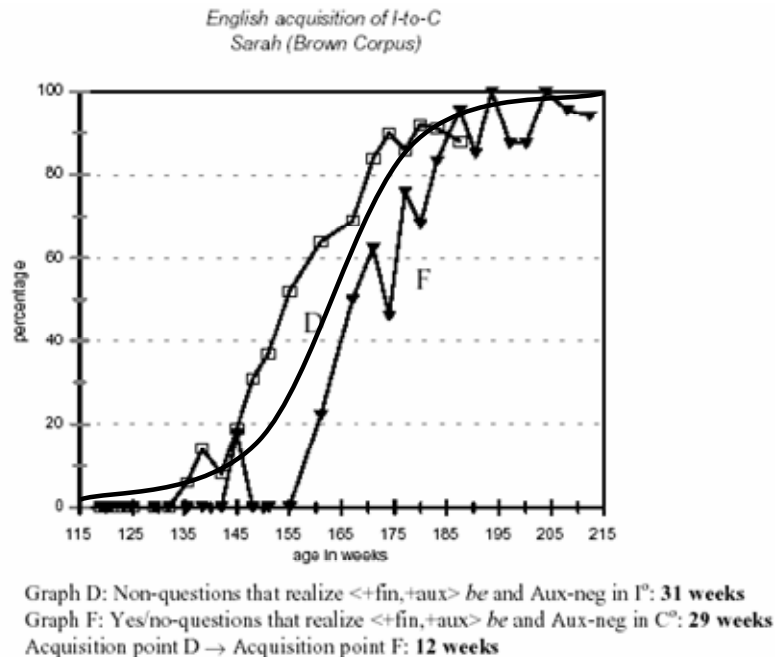


Figure 3 The propagation of an acquisition follows a sigmoid curve

An explanation like the one of Principles and Parameters (a parameter, takes a new value and this determines the application of a new rule) is in a bad position to say why a rule does not apply everywhere at once. Incidentally, this constitutes another argument to abstain from positing rules.

Stochastic approaches explain these transition periods by the coexistence of two stochastic rules and the gradual evolution of the probabilities that weight them. But this explanation does not make a precise causal link between the evolution of stochastic weights and the occurrence-based experiences of the speaker (the language acts).

In an exemplarist and proximalist model as the one defended here on the contrary, the explanation may be more precise if we recall what was just proposed: the occurrence-based licensing of a novel form by a structure mapping with one exemplar – or a few exemplars – already present in the plexus. The analysis process performs mappings between the new form and licensing exemplars. Assume for a new form F1 that the licensing exemplars are P1a, P1b, etc., they are determined proximally depending on the terms that it is possible to recognize in F1. For a form F2, the licensing records P2m, P2n, etc. may accidentally match the P1i but they are most often different, at least at the beginning of the acquisition process. This allows us to understand how, at a given point of the learning process, F1 may take advantage of a new syntactic acquisition whereas F2 cannot yet. Later, P2m for example may have been reanalysed and become aligned on the new syntactic acquisition; this modifies the outcome of the new forms that tend to be licensed by it: a step has been made in the generalization of a usage, the dynamic has progressed a little along the sigmoid curve.

The model lends itself to formalization. Let n , a function of time, be the fraction of the linguistic knowledge that follows a new usage. At a given point of the learning process, the variation of n , that is its derivative function, is proportional to n because, in the proposed schema, only existing exemplars of a construction can license new ones. The derivative of n is thus proportional to n : $dn / dt = k n$. But the only exemplars likely to adopt the new usage are those that have not done it yet. Their number is $(1 - n)$ so the variation of n is also proportional to $(1 - n)$ and the derivative then has the form: $dn / dt = a n (1 - n)$

where a is a factor which is constant with time. The function $n(t)$ itself is obtained by integrating its derivative: $n = \int a n (1 - n) dt$, that is, after integration: $n(t) = 1 / (1 + e^{-at})$

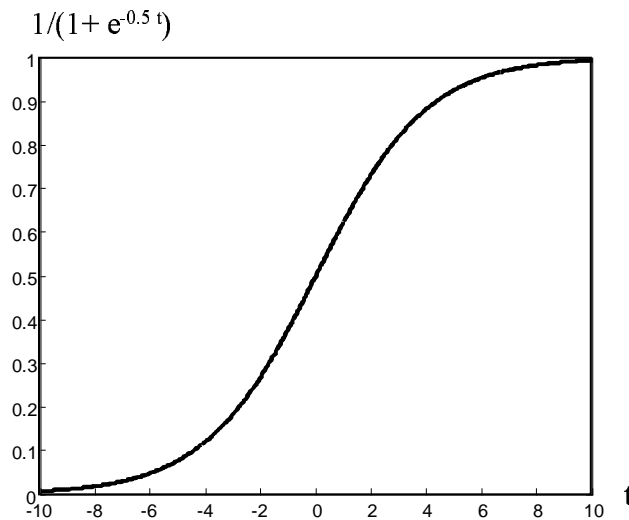


Figure 4 The logistic function

This function is drawn in the figure above, it is the logistic function which governs many social and biological phenomena like the growth of a culture of bacteria, or the propagation of epidemia (when no contrary factor comes to limit their growth). The simple learning model thus predicts that the acquisitions will spread according to the logistic function

Now the logistic function is one of the possible realizations of a sigmoid curve. The prediction of the simple acquisition model is then in accord with the curves which experience shows. Finally, proximalist inscriptions and the simple learning model make the Analogical Speaker a plausible model of linguistic learning.

10 Exemplarism in short

We just reviewed briefly two different exemplarist models: one by Pierrehumbert for phonology, and the Analogical Speaker which I propose for morphological and syntactic productivity. Both make little or no reference to abstractions (in the case of the phonological model, the abstraction 'phoneme' seems to be motivated only by the domain limit of Pierrehumbert's model).

They are efficient with a simpler theoretical apparatus than preceding symbolist models. This simplicity and their not calling on abstractions make them more plausible: it is easier to see how neurons might implement the base functions they posit. An important reason for their efficiency is that they integrate variation, acquisition and diachronic change.

Both contain a notion of proximity which is foreign to all symbolist models and to most connectionist models³², but proximity is to be understood differently in both. In the phonological model, exemplars are close or remote in a space of physical parameters; it is a parametric proximity. In the Analogical Speaker on the contrary, proximity is *de facto* and is supposed to result from usage: it is an inscriptional proximity. The difference is not necessarily a contradiction: it is understandable that phonetics is adherent to physics,

and that, by continuity, this still bears on phonology, but that the proximities active in morphology or syntax are remote from the physics of sound or from that of the vocal tract.

In the late 20th century, linguistic theories tended towards lexicalisation and the exemplarist track prolongs this trend in the following way. Lexicalisation concerned the demotion of abstractions like broad categories and general rules for the benefit of lexical entries, the description of which was complex, and for the benefit of general mechanisms like inheritance and unification. Exemplarism now in turn tends to demote lexical entries; the lexicon itself is touched. There are no longer, or less and less so, monadic decontextualised lexical entries. What are substituted to them are terms the role of which is reduced to manifest the recurrence of some identity. Terms tend to be property-less and this draining is held as indispensable to escape categorical frictions and favour the gradience and flexibility of effects. As a counterpart, the inscriptions where terms are involved 'co-position' them in structural and dynamic microcontexts. It is this recontextualisation which allows us to differently rebuild the productive dynamics. By the same pace, it also allows us to relate in an incremental way, and in a conjecturally causal way, the evolution of the linguistic knowledge with the linguistic acts³³.

We see finally that exemplars, the terms that they involve, and the co-positioning among them, tend to become unessential: they are not constituted out of the contexts and processes which motivate them and set them into operation.

11 Conclusion, from the linguistic subject to the cognitive subject

The reported exemplar-based models are meant to apply to the *factum linguae* and not to a language. A language, specifically an I-language *à la* Chomsky is refused as a hypothesis, along with abstractions. They are also characterised by the option to conjoin the dynamics with a static approach right from the beginning.

The model of the Analogical Speaker, accounts for the recursive productivity which takes place among linguistic forms. Since Lashley³⁴ the cognitive subject is also assumed to present a recursive productivity similar to that of the linguistic subject. This idea has been stressed again several times, still recently by Jackendoff³⁵. This constitutes a first reason that suggests the potential interest to broaden the cognitive subject to the principles adopted in the Analogical Speaker.

A second one is that, research in 'qualitative simulation' results in a combinatorial explosion when it is based on a symbolic approach, notably because it is 'primarist' in the sense defined in section 5 above:

The qualitative simulation algorithms developed to date are problematic as models of human reasoning. Current qualitative simulation algorithms operate via first-principles reasoning over general-purpose axiomatic knowledge. They often produce a huge number of possible behaviors (hundreds or even thousands) even for relatively simple situations. The reason for this is that qualitative simulation, because of the decreased resolution of information about a state, tends to be ambiguous. In a quantitative simulation there is a unique next state, but in qualitative simulations, there can be several next states, corresponding to different transitions that are logically consistent with the resolution of the qualitative state information. Each of these has several next states in turn so their number grows exponentially ... which makes such algorithms seem psychologically implausible, given how easily people reason about everyday physical situations.

A second problem with first-principles qualitative simulation algorithms as models of human commonsense reasoning is that their predictions tend to include a large number of spurious behaviors that logically follow from the low-resolution qualitative descriptions that they use as input but are not in fact physically possible. ... this is not a viable option for modelling the commonsense of the person on the street, who is capable of making reasonable predictions even without such detailed information (Forbus & Gentner 2001 : 35).

Precisely, the Analogical Speaker escapes primarism; it makes it possible for computations to operate at a level of detail just sufficient, contingently depending on the case, by using dynamic analogy and proximity. This constitutes a second reason to apply these themes to cognitive modelling³⁶. It is interesting to note that, in order to circumvent the computational explosion, the same authors also consider analogy with proximal scope as a possible track:

We [Forbus and Gentner 1997] suggest that the solution to this puzzle lies in our use of within-domain analogies (e.g. literal similarity) in commonsense reasoning. We claim that a psychological account of qualitative reasoning should rely heavily on analogical reasoning in addition to reasoning from first principles. Qualitative predictions of behavior can be generated via analogical inference from prior observed behaviors described qualitatively. Prediction based on experience reduce the problems of purely first-principles qualitative reasoning, because they are limited to what one has seen (Forbus & Gentner 2001 : 35,36).

These two remarks seem to me to show why an approach based on exemplars, on proximity, and on a dynamic analogy might be attractive in cognitive science. It would not solve all questions but it would account for the productive explosion. To use the words of Forbus & Gentner (just quoted) the Analogical Speaker may well be said to 'add an analogical reasoning' to one 'based on first principles' but at the expense of a terminology accommodation bearing mainly on 'reasoning'.

The schema is the following. A first mechanism (not in the scope of the Analogical Speaker model) is assumed to establish a first base of meshed dynamic analogies, a first plexus. It does so using low-level perceptual data; in this sense it is primarist; this process is slow, cognitively expensive – because it is primarist – and because of that, the number of its productions is low (hundreds, thousands); it corresponds to the domain of *analogy making* (Gentner typically). This first mechanism is the only one possible and the only one active at the first age of life; with time elapsing, it never really disappears but its comparative importance reduces. A second mechanism (the integrative and dynamic abduction of the Analogical Speaker) takes over. It assumes the availability of the scarce and costly analogies produced by the former and abducts more analogies on the fly, potentially innumerable. It has a low cognitive cost, it is highly productive and it is assumed to be responsible for the productive explosion, first cognitive, then linguistic.

The model of the Analogical Speaker proves this in part in the properly linguistic domain; as a conjecture, it invites to view cognitive acquisition and psychology along these lines. If this track proves workable, it has the potential to account for the productive explosion without being exposed to a combinatorial explosion.

References

Abney, S.

1996 Statistical methods and linguistics. In J. Klavans and P. Resnik (eds.), *The balancing act*, Cambridge, MIT Press.

Chomsky, N.

1957 *Syntactic structures*. Den Haag/New-York: Mouton.

Chomsky, N. and Halle, M.

1968 *The Sound Pattern of English*. New York: Harper and Row.

Chomsky, N.

1984 La connaissance du langage : ses composantes et ses origines. *Communications* 40: 7-24. (Original publication in *Philosophical transactions of the Royal Society of London* 1981.)

Chomsky, N.

1991 *Theorie du gouvernement et du liage, les conférences de Pise*. Paris : Seuil. (Original title: *Lectures on government and binding*. Foris Publications: 1981.)

Chomsky, N. and H. Lasnik

1995 The theory of Principles and Parameters. In Chomsky, N. (ed.), *The minimalist program*. Cambridge: MIT Press.

Engel, P.

1996 *Philosophie et psychologie*. Paris: Gallimard, coll. Folio/Essais.

Forbus, K.

2001 Exploring analogy in the large. In Gentner, D., Kolyoak, K.J. and Konokiv B.N. (eds.), *The analogical mind, perspectives from cognitive science*. Cambridge Massachusetts: MIT Press.

Forbus, K and D. Gentner

1997 Qualitative mental models: Simulations or Memories ? *Proceedings of the 11th international workshop on qualitative reasoning*: 99-104.

Gardner, H.

1985 *Histoire de la révolution cognitive, la nouvelle science de l'esprit*. Paris: Payot. (Original publication 1985.)

Goldinger, S. D.

1996 Words and voices : Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology : Learning, Memory, and Cognition* 22: 1166-83.

Habert, B.

2003 Régler les règles. *Traitement Automatique des Langues*: 43-3.

Harris, Z. S.

1991 *A theory of language and information. A mathematical approach*. Oxford: Clarendon Press.

Hintzman, D. L.

1986 Schema abstraction' in a multiple-trace memory model. *Psychological Review* 93: 328-38.

Householder, F. W.

1971 *Linguistic speculations*. Cambridge Massachusetts: Cambridge University Press.

Hummel, J. E. and K.J. Holyoak

1997 Distributed representations of structure : a theory of analogical access and mapping. *Psychological Review* 104 (3): 427-66.

Itkonen, E. and J. Haukioja

1997 A rehabilitation of analogy in syntax (and elsewhere). In A. Kertesz (ed.), *Metalinguistik im Wandel: die kognitive Wende in Wissenschaftstheorie und Linguistik*. Frankfurt a/M: Peter Lang.

Jackendoff, R.

2002 *Foundations of language*, Oxford: Oxford University Press.

Johnson, K.

- 1996 Speech perception without speaker normalization. In Johnson, K. & Mullemix (eds.), *Talker Variability in Speech Processing*. San Diego: Academic Press.
- Lavie, R. J.
- 2003 *The Analogical Speaker or grammar put to its place*. Doctoral thesis. Nanterre: Université de Paris 10.
- Lavie, R. J.
- 2004b Systemic productivity must complement structural productivity. *Proceedings of the Xth Congress of Cognitive Linguistics, Braga, Portugal, July 2003*.
- Livet, P.
- 1995 Connexionnisme et fonctionnalisme. *Intellectica* 21.
- MacWhiney, B.
- 2000 Connectionism and language learning. In Barlow M. and S. Kemmer (eds.), *Usage-based models of language*. Stanford: CSLI Publications.
- Manning, C. D.
- 2002 Probabilistic syntax. In Bod, Hay and Jannedy (eds.), *Probabilistic linguistics*. Cambridge: MIT Press.
- Millikan, R. G.
- 2000 *On clear and confused ideas*. Cambridge USA: The Press Syndicate of the University of Cambridge.
- Minsky, M. L.
- 1985 *The society of mind*. New York/London: Touchstone, Simon & Schuster.
- Mulder, Jan W.F. and P. Rastall
- 2005 *Ontological questions in linguistics*. Munich: Lincom GmbH.
- Pereira, F.
- 2000 Formal grammar and information theory : together again ? *Philosophical Transactions of the Royal Society* (358): 1239-1253.
- Pierrehumbert J. B.
- 2001 Exemplar dynamics : word frequency, lenition and contrast. In Bybee, J. & P. Hopper (eds.), *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins.
- Ploux, S. and B. Victori
- 1998 Construction d'espaces sémantiques à l'aide de dictionnaires de synonymes. *Traitement automatique des langues* 39 (1):161-182.
- Putnam, H.
- 1960 Minds and machines. In S. Hood (ed.), *Dimensions of mind*. New York: New York University Press.
- Skousen, R., D. Lonsdale, and D. B. Parkinson (eds.)
- 2002 *An exemplar-based approach to language*. Amsterdam/Philadelphia: John Benjamins.
- Smolensky, P.

1999 Grammar-based connectionist approaches to language. *Cognitive Science*: 23-4.
van Kampen, J. and A. Evers
2000 *E-language, I-language and the Order of Parameter Setting*,
<http://www.let.uu.nl/~Jacqueline.vanKampen/personal/downloadables/evers-kampen-Syntax.pdf>.

¹ In the literature so far the words used are *exemplar-based models*, *exemplar-based theories*. To collectively refer to such approaches, I shall use *exemplarism* not because they ought to be taken as examples but because they are exemplar-based.

² Exemplar-based models share with connectionism the demotion of abstractions and some amount of parallelism; they are models of an individual speaker (not of a language); they involve a vast population of elements; they are dynamic and feature competition; each element in them may be useful but none is indispensable. In the yielding of results, they combine various viewpoints and produce gradual effects. In short, exemplar-based models have the qualities attributed by Livet (1995) to connexionist models: "*a local compositionality and a limited systematicity*".

³ I understand 'constraints' in the sense of this word in Optimality Theory in linguistics. Cf. Smolensky 1999 for a brief introduction contrasted with generative grammars.

⁴ For their help in various respects, I thank B. Laks, J. Lassègue, V. Rosenthal, M. Toussaint, Y.M. Visetti and two anonymous readers.

⁵ In phonetics, F0 is the fundamental frequency of a signal, F1 its first formant, etc.

⁶ In the context of this paper, such *labels* are phonemes exclusively (/e/, /i/, etc.). In the framework of exemplar theory, which generally tends to demote the categorial viewpoint, one may be intrigued by a reference to a phonematic category system. Below, I shall state what reading should be made, in my view, of their utilisation by Pierrehumbert.

⁷ In linguistics, 'lenition' is the diachronic process whereby a phoneme ends up not being pronounced.

⁸ Thus far, we were considering a constant phonological 'system', even if the parametric grounding of phonemes might evolve in time. A neutralisation now modifies the system itself.

⁹ Even a prototype theory enhanced with family resemblance or augmented with a stochastic complement.

¹⁰ But this assumption is not inherent in exemplar theory: exemplar theory just requires a complementary assumption about how exemplars are affected by time and several schemas remain possible.

¹¹ Millikan 2000:42.

¹² Lavie 2003.

¹³ In 'static' we must comprehend also generative grammars because the 'process' of derivation (then of transformation) which they encompass, is deployed in a space which is not that of the linguistic dynamics. Generativism has little to say about the linguistic dynamics. For a recent statement on the subject, cf. Jackendoff (2002 : 57).

¹⁴ Harris 1991, Abney 1996, Pereira 2000, Manning 2002, Habert 2003, etc.

¹⁵ For details about stochastic complements to categorial theories, cf. Lavie 2003:238.

¹⁶ By 'dynamics', I understand a) the reception of the linguistic material and its interpretation in a situation, b) its emission still in a situation, then, as an effect of the latter ones, c) the acquisition of a linguistic know-how, and finally d) linguistic change.

¹⁷ This state of affairs was prevailing in the 1950s, and Chomsky availed himself of that to disqualify analogy. He approached the question taking another course, breaking in this way with recent history, and rooting himself back in the French 17th century.

¹⁸ This enterprise also connects with the more recent comeback of analogy in the last decades of the 20th century, in psychology (Gentner, Holyoak), and in cognitive science. Analogy also comes back, less predominantly in linguistics: Lakov, Itkonen (1997).

¹⁹ The example is an excerpt of the English plexus which will support the case studied in section 6 below. Plexii with some representativity were written for a speaker of French (2000 terms), for a speaker of English (800 terms), and smaller ones for speakers of other languages.

²⁰ Inscriptions arranged into structure mappings in this way are called 'paradigms' in the jargon of the model, which extends a little the acceptance of the word in structural linguistics.

²¹ This clause is simplified for pedagogy. Applied so strictly, it would make the model under-productive; the model's reality is more shaded but this is beyond the scope of this article. The overall principle is the one exposed though.

²² The example is provided in French as the morphology of English is less varied and lends itself less to show simply what I mean with 'systemic analogy'.

²³ Details in Lavie 2003:139 or Lavie 2004.

²⁴ More recently, in Chomsky's Minimalist Programme, it supposes a rich apparatus of syntactic features.

²⁵ Lavie 2003:112.

²⁶ 'Structural productivity' is that which motivates new assemblies.

²⁷ I call 'systemic productivity' the ability to place linguistic forms in pluridimensional paradigms and to collect meaning effects attached to such placements.

²⁸ Details of this argument in Lavie 2004.

²⁹ About this level, Engel (1996) uses the word "normative". For example : *Frege reproaches the psychologists with confusing two meanings of the word "law" when they equate logical laws with psychological laws: the sense "normative" and the sense "descriptive". Even assuming, as he does, that logical laws are normative laws, Frege still confuses two meanings of the word "norm": a sense in which a norm describes the laws of an intelligible universe, and one in which a norm prescribes to individuals to follow a certain rule.* Engel 1996, p. 120.

³⁰ We talked before (section 3 above) of the propositions for improving symbolist grammatical theories by adding to them a stochastic complement.

³¹ The argument is recalled and summarized in Gardner 1987/1993, p. 45.

³² However, connectionist models which draw on *self-organizing* maps, e.g. McWhinney (2000), do have a notion of proximity; some models of lexical semantics also have, e.g. Ploux and Victorri (1998).

³³ In the ongoing debate about ontology in linguistics, see for example Mulder & Rastall (2005), the Analogical Speaker takes a specific stance : it refutes the object-property opposition and fosters instead the vacuity of the terms (they are property-less) and their analogical copositionings.

³⁴ 1948, the Hinton symposium; this is recalled by Gardner (1987/1983 : 23-26).

³⁵ Jackendoff, conférence à la MSH, Paris, printemps 2003.

³⁶ Incidentally, the *spurious behaviors* reported by Forbus recall the dilemma met in natural language processing by the designer of parsers when he adopts a categorical frame – that is, always so far. Either he works with a restricted set of categories and then the rate of successful analyses is low; or he refines his category set and then he has to face an explosion of induced syntactic ambiguities. The 'desambiguation' overhead may become prohibitive, and the permanent complaint we hear is that it is impossible to reconcile accurateness and efficiency. This must be regarded as one more symptom tending to show the weakness of categorial theories.