Kristian Tylén, Riccardo Fusaroli, Peer F. Bundgaard and Svend Østergaard

Making sense together: A dynamical account of linguistic meaning-making*

Abstract: How is linguistic communication possible? How do we come to share the same meanings of words and utterances? One classical position holds that human beings share a transcendental “platonic” ideality independent of individual cognition and language use (Frege 1948). Another stresses immanent linguistic relations (Saussure 1959), and yet another basic embodied structures as the ground for invariant aspects of meaning (Lakoff and Johnson 1999). Here we propose an alternative account in which the possibility for sharing meaning is motivated by four sources of structural stability: 1) the physical constraints and affordances of our surrounding material environment, 2) biological constraints of our human bodies, 3) social normative constraints of culture and society, and 4) the local history of social interactions. These structures and constraints interact in dynamical ways in actual language usage situations: local dialogical and social dynamics motivate and stabilize the profiling of a conceptual space already highly structured by our shared biology, culture, and environment. We will substantiate this perspective with reference to recent studies in experimental pragmatics and semiotics in which participants interact linguistically to solve cooperative tasks. Three main cases will be considered: The dynamic grounding of linguistic categories, the construction of conceptual models to relate entities in a scene, and the construction of shared conceptual scales for assessing and appraising subjective experiences.

Keywords: meaning-making; categorization; conceptual models; affordances; social dynamics

Kristian Tylén: Aarhus University. E-mail: Kristian@cfin.dk
Riccardo Fusaroli: Aarhus University. E-mail: fusaroli@gmail.com
Peer Bundgaard: Aarhus University. E-mail: sempb@hum.au.dk
Svend Østergaard: University of Aarhus. E-mail: semsvend@hum.au.dk

* The authors would like to thank the Velux Foundation, the Danish Council for Independent Research – Humanities, and the EUROCORES project: Digging the Roots For Understanding (DRUST).
1 On the public nature of meaning

It has been argued that for linguistic communication to be possible, interlocutors must share the same public language (Wittgenstein 1958:§246): this entails not only to share a set of linguistic practices for designating referent objects. Rather, successful communication depends on the degree to which people share a common set of invariant “senses” that mediate between spoken or written word forms and their reference matter – what is often referred to as “concepts” or “meanings.” These invariant aspects of meaning are thought to transcend locally contextualized instantiations and the episodic and subjective psychology of individual language-users. But what is the ground for this invariance of meaning? How is it that we can reliably assume that our interlocutors will understand what we say in this or that context?

1.1 Grounding invariant patterns of meaning

Often-cited classical positions hold that meanings must already “be there” for people to point to them in communicative acts. Meaning entities are thus claimed to have an “ideal” mode of existence independent of individual language users. One solution to this problem has been to claim that meanings belong to a Platonic realm of ideas or intelligible forms of which the material objects are only an instantiation. In this sense, the intelligibility of a given material object stems from the idea of which it is only an incomplete copy (Frege 1950, 1960). Another position, for instance found in traditional interpretations of the works of Ferdinand de Saussure, explains the invariant properties of meaning in terms of language-internal differential, invariant relations: “Language is a system of interdependent terms in which the value of each term results solely from the simultaneous presence of the others” (Saussure 1959: 114). The meaning of a word is thus defined by the network of immanent linguistic relations and does not depend on anything external to language itself (i.e., extra-linguistic elements, such as references or local use)1. Both of these positions tend to reduce the possibility of communication and understanding to reliance on pre-existing meaning configurations that are located beyond the individual. This leads to at least two widely acknowledged

---

1 In fact, Saussure defines linguistic value as being determined at the same time by linguistic and extra-linguistic relations:

two features are necessary for the existence of any value. To determine the value of a five franc coin, for instance, what must be known is: i) that the coin can be exchanged for a certain quantity of something different, e.g., bread, and ii) that its value can be compared with another value in the same system. (Saussure 1959: 160).
problems: the objectification and unconstructive discussions of the “locus of meaning” (Lakoff and Johnson 1980; Putnam 1973, 1975) and the “symbol grounding problem” including questions such as “how does meaning emerge? How do we bootstrap our understanding of symbols?” (Harnad 1990; Steels 2008).

In addition, there is a tendency in classical theoretical approaches to generalize from rather peripheral domains of meaning-making such as science, logic, and mathematics and thus lose sight of the actual everyday functional and intersubjective contexts in which linguistic meanings do their job; as means of social coordination, interaction, and communication (Tylén et al. 2010). However, and this is the crucial point: while it is decisive for some domains of knowledge and communication, for example, mathematics, that we mean exactly the same thing when we talk about triangles or establish the concept of natural numbers, this might not be the case in the vast majority of everyday communicative situations. Rather than sharing the exact same meanings, we use linguistic utterances to disambiguate a shared context (Collier and Talmont-Kaminski 2005). Indeed communication most often happens in richly structured and value-saturated contexts, where only quite minimal efforts are needed to reach common understanding (Croft 2010a).

A more recent approach to the nature of meaning referred to as “embodiment theory” stresses that linguistic meaning is indeed shaped by 1) our psychophysical makeup, and 2) our bodily engagements with the material environment (Lakoff and Johnson 1999: 16). Meaning is thus intrinsically linked to our (individual) bodies, brains, and minds. Since physical regularities exist, and since we share the same biological architecture and thus the same constraints on our experiential engagements with the environment, we come to categorize and conceptualize the world in similar ways in the form of “cognitive schemas” (Lakoff 1987) or “image-schemas” (Johnson 1987). Such schemas are abstract skeleton representations that are claimed to be universally available as a motivational source for language structure (Talmy 2000). Though we are generally sympathetic to such approaches, some of its manifestations still share reminiscences of the problems touched upon above. Meaning is thus not grounded in online functional contexts of language use, but rather in solitary individuals early experiences of having a body and moving in space (Lakoff 1987; Lakoff and Johnson 1980, 1999; for a related critique, see Thibault 2004), which makes the approach incapable of addressing issues of, e.g., language change.

2 The fallacy consists, thus, in saying that if exact meaning entities don’t exist in everyday cognition, then mathematics is a fiction, as Frege had it. The domains are different and the requirements are different: absolute exactitude is an issue in mathematical or scientific knowledge, not in everyday cognition and communication.
While we do find pronounced regularities in public linguistic meaning, we question whether these pertain to transcendental ideal properties. Typological studies have made it increasingly clear that practically no universal principle – including the existence of word classes like verbs, or even the core grammatical machinery of recursion, constituency, and grammatical relations – applies to every single language in the world (Croft 2001, 2003, 2010b; Evans and Levinson 2009). Other studies point to the fact that not even within the same language do all speakers converge on essentially the same grammar, and profound, systematic differences in individual speakers are found for instance in relation to socio-economic conditions (Dabrowska and Street 2006; Street and Dabrowska 2010). These observations are supported at a local interaction scale by a number of recent studies in experimental pragmatics revealing that even when participants use language to solve tasks in very constrained experimental environments, they show profound variation in the linguistic strategies they evolve and employ (Croft 2010a; Garrod and Anderson 1987; Garrod and Doherty 1994; Healey and Mills 2006). Together, these findings suggest that universally invariant aspects of meaning may in fact neither be found in inherent formal language structure (whether on the level of words, constructions or discourse), nor in any specific embodied conceptualization underlying such language structures.3

In the following, we will elaborate on an alternative approach stressing how multiple sources of structural stability are selected and realized in the dynamical construal of meaning in concrete language-usage situations.

1.2 Towards a dynamic grounding of meaning

... the grounding of symbols is not a once and for all affair, but a matter of progressive and continuous coordination.
– Steels (2008: 8)

Linguistic utterances unfold in pragmatic contexts where interlocutors actively pursue communicative goals such as the coordination of actions, sharing of perspectives, knowledge, fictions or past experiences, etc. Such contexts encompass the material and social setting, the interlocutors’ local history of interaction and their communicative goals (Clark and Brennan 1991). They are highly structured and semantically potentiated situation spaces and thus constitute a rich and reliable semiotic resource for joint meaning-making (Goodwin 2000, 2007). While

3 At least they are not determined by embodied conceptualizations as seems to be a common conception in some manifestations of Cognitive Linguistics.
some properties of the situational common ground will have an ad hoc character, there are other more stable structural properties that we will argue play a crucial scaffolding role in meaning-making. We identify four such main sources of structural stability:

- Our biological makeup constrains our sensorimotor engagements with the world. Though subject to great plasticity, there are limits to our perceptual sensitivity (e.g., the range and resolution of our perception, etc.) and our motor engagements with the environment that naturally constrain the way we can conceive of the world.

- Our physical/material world has stable structure affording (but not determining) certain categorical and conceptual distinctions rather than others. This physical world encompasses both “natural” and “enculturated” properties and objects since a long history of human intervention has profoundly shaped our material world into a hybrid of natural objects and artifacts.

- Our immersion in specific socio-cultural and linguistic practices puts profound normative pressure on the way we perceptually attend to and realize the multiple structural affordances of the material environment. This gives local lingual-cultural variations to the structural affordances realized in different languages.

- The local history of successful communicative interactions constitutes another set of social affordances that, through processes of reciprocal alignment of perspectives and profilings, scaffold the stabilization of linguistic meaning.

The idea offered here is that stabilities in communicative meaning are motivated by a multiplicity of structural invariance patterns grounded in our shared biology and material and socio-cultural environments. These structures scaffold and constrain online human meaning-making, but – most importantly – they do not determine meaning. Rather, the concrete linguistic realizations of the structural potentialities like the ones outlined above critically depend on the local history of dialogical interactions in which interlocutors align their linguistic construals of a situation in pursuit of communicational needs. And it is in these dialogical realizations and profilings of situated structural affordances that meaning-making unfolds.

4 A possible fifth source of stability – that will not be treated here – is constituted by the “form-side” of communicative signs. Rather than purely conventional, arbitrary relations, the meaning of words and sentences can also be motivated by the iconic and diagrammatic intuitions we ascribe to material phonetic/prosodic or graphical patterns of sign tokens (cf., e.g., Fay et al. 2004; Galantucci et al. 2010).
Any situation (including even the most constrained situational space such as in thoroughly controlled experimental environments) can be linguistically construed in multiple ways, realizing one or another of the situation-inherent structural affordances or perspectives. When referring to a situation, speakers thus have the option of various possible constructions. Consider a simple example where a human participant is involved in an event with an unintended outcome. The situation can be structurally construed in three different ways: 1) with the experiencer as the subject: “he crashed into a rock,” 2) with another participant encoding the subject: “his bike hit a rock” (foregrounding the bike), or 3) with an existential: “there was a rock in the way so the bike fell over” (foregrounding the obstacle). This plurality of possible profiles is a general phenomenon. The same situation can afford multiple construals. But the repertoire is far from infinite or random. In this example, physical properties of the situation including the order of events, spatial and causal organization, agent-patient relations, etc., constrain the conceptual organization and thereby the way interlocutors make sense of the situation linguistically. The choice of one construction over another may be motivated by several factors. Michael Tomasello has thus argued consistently for the perspectival nature of language (1999, 2003): different constructions represent different alternative perspectives on the same situation. The perspectival choice can have various local pragmatic motivations, either in relation to the preceding discourse, or intended rhetoric effects. But there can also be language-external motivations that make a particular constituent part perceptually more salient and thus afford a certain linguistic profiling. Besides, there are lingua-cultural macro-variations in the way similar situations are profiled linguistically. A given language can have normative preferences for some construals rather than others (Berman et al. 1994). Interestingly, though, these variances are again not random, but can be interpreted as complementary profilings of the same “conceptual space” (Croft 2001, 2010b).5

More crucially for the argument here, there are also profound stochastic variations in the way interlocutors (of the same mother tongue) linguistically realize a referential scene that cannot be explained by language-specific normativity or perspectival saliencies. In the course of situated dialogical interaction, interlocutors may from time to time favor specific ways of labeling objects or construing situations that work for that particular interaction in that particular context. Any such locally negotiated practices will – to the extent they are successful in achiev-

5 When translation and cultural understanding is in fact possible, it is because different languages realize different aspects of the same, fundamentally invariant structural affordances of the types listed above.
ing their communicative goals – have the potential for normative stabilization and spread through repeated interactions (Garrod and Doherty 1994).

1.3 The dynamics of dialogue

Central to the general theory of meaning-making put forth here, is the role of situated dialogical dynamics. Several elements are important for a full description of the local dynamics of dialogue; here we will however limit ourselves to a few.

An important feature is often referred to as “common ground,” i.e., the degree to which interlocutors share the same referential space, for instance, due to their common cultural or professional background, etc. However, common ground is not a static thing: there is a constant update of common ground during conversation (as demonstrated in, e.g., Clark and Brennan 1991) that facilitates communicative consensus with a minimum of linguistic effort.

Another important aspect of dialogue is the propensity of interlocutors to spontaneously align their vocabulary, syntax, etc. This has, for instance, been demonstrated in a study by Branigan et al. (2000), where pairs of participants had to describe pictures to each other. One of the participants was a confederate of the experimenter who would intentionally vary the syntactic structure of her image descriptions. This was found to spontaneously prime the other participants’ use of similar syntax in subsequent descriptions. This is an unconscious mechanism thought to be involved in the propagation of stable linguistic forms in a community. In Pickering and Garrod (2004) and Garrod and Pickering (2009) this kind of imitation is argued to take place on several levels: postures, speech rate, choice of words, and choice of syntactic construction leading to an alignment of a shared “conceptual model” of the topic of the conversation. Interestingly, however, the imitator has to be a true participant of the dialogue. In Branigan et al. (2007) it is demonstrated that overhearers do not imitate syntactic forms to the same extent as the addressee. Participatory dialogical practices thus seem to be crucial for the stabilization, change, and spread of linguistic patterns in a community.

While the scaffolding structural affordances listed in the section above constitute the source of invariance in meaning, local social dynamics is thus the driving force in the continuous rich linguistic realizations of these potentialities and thus constitute the core of language evolution and change (Loreto and Steels 2007; Steels 2008). A number of recent empirical studies have investigated various aspects of the interaction between structural potentialities and social dynamics.
2 Exploring collective meaning-making

In the subsequent sections, we will review a number of recent experimental studies of various kinds of meaning-making practices. Common to them is that they pertain to situations where interlocutors have to communicate in order to coordinate the joint solution to a problem: the interlocutors need to construct a shared operative representational layer that enables joint action. In these contexts, language is thus considered the tool by which interacting minds coordinate, structure, and constrain non-linguistic interactive cognition (Tylén et al. 2010) via alignment and negotiation of the affordances of the situated interactive practice. While the situation is constrained by the need to solve the task at hand, the linguistic and semiotic exchanges tend to display at the same time a limited set of common strategies and wide variability in developing and unfolding them.

We will focus especially on three quite different domains of meaning-making varying in complexity and scope. The first concerns the dynamic grounding of linguistic categories, that is, processes underlying how interlocutors conjointly parse a scene in discrete linguistic categories, for instance, when motivated by salient discontinuities. The second is about how interlocutors construe conceptual models to relate entities in a scene, for instance, for the purpose of navigation. The third concerns how interlocutors construe shared conceptual scales for assessing and appraising subjective experiences.

2.1 The dialogical route to linguistic categories

Language is uncontroversially crucial in expressing categories – that is, ways of recognizing, differentiating, and grouping objects and concepts. The role of language in the development and unfolding of categorization is however still under investigation. One of the most influential theories is prototype theory (Rosch 1975, 1999), defining categories as structures of features that exemplars can display in minor or major degree. Often prototype theory has been associated with an embodied view of cognition (Barsalou 2008; Lakoff 1987): our sensorimotor couplings with the world ground our ability to think and talk about it. The most salient interactive features define the structure of prototypes. However, what is consistently underthematized in these approaches is the role of socio-cultural experience in structuring such prototypes, in particular in relation to language-use and social interactions.

One way to study the impact of social dynamics on categorization processes is to compare how speakers of different languages diverge or converge in their biases towards linguistic realization of specific structural affordances and constraints of a scene. To investigate such issues, Malt and colleagues (2008, 2003)
developed two experiments, one related to gait categorization and the other to artifact categorization.

In the first experiment considered here, Malt and colleagues compare labels for different paces of locomotion in English, Japanese, Spanish, and Dutch (Malt et al. 2008). Native speakers of these languages had to describe locomotion at different slopes and speed performed by a person on a treadmill. For this particular type of stimuli, bio-physical affordances turn out to play a crucial role: human gaits all over the world are highly constrained by the bio-physical dynamics of motion (Collins and Stewart 1993), giving salient discontinuities between walking and running gaits as speed increases, rather than gradual shifts through intermediate versions (Schöner et al. 1990). But there are also normative linguistic constraints at work: In English and Dutch, verbs tend to express manner of motion, while in Spanish, they express direction of motion, and in Japanese, verbs tend to express directional path or path plus ground.

The comparison of languages reveals a highly constrained semantic space: names for two common human gaits (“walk” and “run”) follow the same categorical pattern across the different languages and the best examples of the gait terms are highly comparable across languages. There is thus a critical point in the human physiology of gait separating “walk” from “run,” that seem to be of universal significance, however, cross-linguistic comparisons show great variability in how the two areas are further subdivided. The author thus conclude: “These findings support the proposal that structure in the world provides constraints on how category labels are assigned . . . Where strong structure exists, broad categories may tend to be shared across languages” (Malt et al. 2008: 239).

When this approach is applied to cultural artifacts, containers in particular, however, the bio-physical affordances seem more complex and less salient allowing for more radical cultural variability (Malt et al. 2003). In the experiment, English, Spanish, and Chinese speakers were asked to name sixty containers. The results point to pronounced cultural-linguistic differences in how categorical boundaries are drawn in the three languages: Though some boundaries and prototypes were shared, nesting (several categories in one language falling into a single category of another language) and cross-cutting categories (parts of two or more categories in one language constitute a single category in another language, etc.) were quite common. The findings suggest that in the case of cultural artifacts like containers, multiple factors interact in motivating category boundaries. Some of these are physical attractors in terms of salient discontinuities of the stimulus space related to shape, material, type of closing, etc., but in addition, the authors speculate that cultural factors (such as the relative needs to communicate about distinctions within a domain), linguistic factors (e.g., extensions of already existing categories to new cases, morphological motivations, etc.), and
more stochastic variation in the selection of structural potentials all contribute to
the formation of categories in a linguistic community.

In the case of containers, we thus see a multifaceted conceptual space with
several dimensions of affordances (physical, cultural, and linguistic) that can be
realized in category boundaries. However, despite these scaffolding potentials,
the concrete carving of the categorical space is an open and indeterminate pro-
cess that depends on a history of situated social interactions and thus resulting in
the kind of cultural variability explored by Malt and colleagues.

In order to more closely investigate the social dynamics involved in the
formation of categories, a number of new approaches have emerged using
software-modeled algorithms. These can be used to simulate processes of multi-
generational language evolution and change. Among the strengths of such ap-
proaches are the tightly controlled inclusion and manipulation of a multiplicity
of factors like the ones considered above. The settings used to explore category
formations are various kinds of “naming games” (Baronchelli et al. 2005; Puglisi
et al. 2008; Steels 1995; Steels and Vogt 1997). The crucial question that these
simulation games attempt to solve is how shared sets of linguistic categories
stabilize at a global level without any central coordination. The game employed
by Baronchelli et al. (2005) is based on the need of a population of “agents” to
develop and share a common categorization of their environment. Language,
operationalized as a shared set of conventions for mapping meanings to expres-
sions, is the tool employed to achieve that goal. During the game agents interact
in pairs: a hearer and a speaker. The speaker selects an aspect of the environment
– a referent – and utters a linguistic label. If the speaker does not already know
the label for the object, then a new name is invented. If the hearer recognizes the
label, the communication is successful. In contrast, if the hearer does not possess
a match, the trial is a failure, but the hearer now inserts the new label in its mem-
ory. The study primarily focuses on naming coordination: there is a given amount
of pre-defined referents, and the factor that is being manipulated is the strategy
that speakers use to choose between existing names: either a 1) random strategy,
2) using the first name in the memorized repertoire, 3) using the last successfully
used label. The results show that, even with such a minimal principle for com-
munication, a global structure quickly emerges, with the last strategy demanding
fewer interactions. Baronchelli and colleagues thus demonstrate that once con-
ceptualization, communicational channels, and minimal forms of social interac-
tion are in place, shared language structure is simply a consequence of repeated
interactions.

In a refinement of the latter approach, Puglisi et al. (2008) investigate the
construction and negotiation of categories and their boundaries via the selection
of relevant aspects of the environment. The simulated agents are to solve a color
categorization task carving up a continuous, one-dimensional color space. Local environmental effects are simulated by manipulating the color frequencies of the stimulus space (e.g., making the agents categorize the color distributions extracted from images of a snow landscape or a colorful parrot). Besides, the agents’ perceptual engagements with the stimuli are constrained by a simulated biological resolution threshold (the minimal distance between referents that can be discriminated). The question is thus slightly more sophisticated: not only how linguistic coordination on given referents unfolds, but the dynamics through which the continuum of perceivable referents in the world is organized in a small and finite number of stable linguistic categories, even if an infinite number of other categorizations would potentially be possible.

The simulation exhibits the same behavior as in the previous study. After a period of chaos, there is a transition to a common linguistic structure with a certain amount of failures due to unaligned categories and ambiguities. The simulated agents thus converge on a limited number of linguistic categories – usually around ten – though different runs of the game yield different results, due to the indeterminacy of the discrimination task.

Such results seem to reflect real world color category formations and thus conform to the World Color Survey (Baronchelli et al. 2010). Color categorization varies between languages; however, most languages have six to twelve linguistic labels (Berlin and Kay 1969; Lindsey and Brown 2006), while the speakers can still perceptually discriminate between a much higher number of color categories.6

The picture that emerges from these experiments and simulations is that categorization is a complex process in which situated semiotic dynamics interact with bio-physical saliencies (e.g., the structural properties of human locomotion, as well as the biological capacity of discriminating colors), culturally normative constraints (as in the inherited preferences for construals and names for certain locomotions and containers, etc.), and the local history of the interaction itself with its potential stochastic variations.

2.2 Aligning conceptual models

While in the case of categorization we deal with quite simple form/meaning pairings (the establishment of a correlation between labels and reference entities), we

6 Besides, the same basic dynamics has been found in how (living human) Internet users attribute keywords to links, pictures, resources, and movies in popular websites like Delicious.com, Flickr.com, BibSonomy.org and Imdb.com, although the dominant categories seem generally more open to changes over time (Cattuto et al. 2009).
will now address situations in which interlocutors have to refer to more complex relational structure of scenes. When referring to relations in the world, it is not enough to agree on suitable labels to entities; interlocutors have to specify salient aspects of the scenery for referencing. For instance, if we want to locate an entity — say a cat — we can say something like “the cat is lying under the table,” specifying the surface of the table as the salient aspect of the scene to establish the relation. Sometimes, however, speakers in dialog are confronted with novel situations for which there are no obvious conventional way of establishing the relational reference. Through linguistic interaction, they thus have to converge on a shared system of reference. Such cases are addressed in a series of experiments briefly considered below.

In Galantucci (2005, 2009), it is not the specifics of spoken conversation that is the topic, but rather the development of a communication system in general. Pairs of participants have to find each other in a virtual environment consisting of four rooms, each marked with an icon. They can only see the room in which they themselves are located, but they can communicate to each other through a special writing pad. A special feature of the pad (a constant vertical drift) prevents the use of any conventional symbols: letters, numbers, and drawings are distorted so that they cannot be recognized. In other words, to solve the experimental task participants have to create a language (system of graphical signs) from scratch.

The experiment seems to reflect general principles underlying ordinary language dynamics (cf. Garrod and Anderson 1987). Most importantly, the participants rapidly develop communication systems that make them significantly better in solving the task than if they navigate without any form of shared sign system. Besides, the novel sign systems are based on participants’ selection of salient features of the virtual environment as “models of reference” similar to natural language referencing. In some cases, sign systems are based on the icons marking the rooms (for instance, three dots might mark the vertices of the triangle pictured in one of the rooms), while in other cases dyads use the spatial layout of the rooms as basis for their sign systems (for instance, a vertical line at the left side of the pad will show that the participant is in one of the left most rooms). Finally, some pairs of participants converge on using symbolic enumeration; one dot mark the rightmost room at the top, two dots the rightmost room at the bottom, etc. The emergence and use of novel signs is found to be tightly integrated with behavioral coordination. Participants are thus not only reliant on sign exchange but likewise on local stabilization of behavioral practices. For instance, telling that you are in the left part of the maze is enough to secure a winning outcome of the game if participants share implicit practices on who moves and who stays. This observation ties up to the grounding hypothesis of language (Clark
The communicative practices are refined to the minimally necessary given the common ground. This is also referred to as the underspecification of language (cf., e.g., Radden et al. 2007).

Another important feature of the game, analyzed in Galantucci (2009), is the fact that successful coordination in the game depends on implicit online alignment dynamics of the kind proposed by Pickering and Garrod (2004: see section 1.3 above). Between trials, participants have the opportunity to negotiate, for instance, making a sign and at the same time moving towards the icon would naturally mean that the sign refers to the icon. But in fact this kind of explicit negotiation (meta-communication) is negatively correlated with performance.

The findings in the previous experiment are confirmed in a more language-oriented study presented in Garrod and Anderson (1987) and Garrod and Doherty (1995). In these studies, pairs of participants navigate a maze consisting of boxes connected by paths. Again the task is for the participants to end up in the same box, but the game is construed so that to get around each participant has to rely on information from the other (for details see Garrod and Anderson 1987). The participants are observed to use at least four ways of conceptualizing the maze: Some dyads use a “line-model” conceiving the maze as consisting of lines and rows, and thus use expressions such as “second box in the third row” to identify a node. Others use a matrix model similar to the one used in chess, so that “c2” would refer to the same node as the formulation above. Furthermore, some participants use a path-model, navigating the partner to a specific node by naming a path leading to it. Finally, there was some use of a figural model: for instance if there is a T-shape protruding from the maze, they would refer to a box as located at the T-junction, etc.

The findings resemble the ones reported by Galantucci (2005). All the pairs converged on a preferred conceptual model. The model could change and develop during the games, but it did not lead to divergence between the speakers. The choice of a model was to some extent determined by saliency. For instance, if the maze was regular, e.g., six by six, it rather afforded the line or the matrix model, whereas if it was very irregular the figural model was sometimes used.7

The main principle for establishing a common form/meaning pairing was found to be the “principle of precedence” (Garrod and Doherty 1994) or the “output-input principle” (Garrod and Anderson 1987), resembling the findings by, e.g., Baronchelli and colleagues (Baronchelli et al. 2005): A speaker attempts to match as closely as possibly the options he used to interpret the last relevant

7 However, the figure model was not very stable in the sense that if it was used in the first game there was only a low probability (compared to the alternative models) that it was used again in the subsequent game.
utterance from the dialogue partner. Besides, as found by Galantucci (2009), there was no positive effect of explicit negotiation; in fact explicit negotiation was very rare.

To sum up some basic points concerning the interactions dynamics found in these studies, interlocutors tend to converge on a common conceptual model of the situation they are navigating. This convergence is determined by implicit online negotiation and alignment. The main mechanisms involved are selection of salient aspects of the situation and reliance on the local history of the interaction (the principle of precedence).

### 2.2.1 Propagation in the community

The studies considered above address how isolated pairs of interlocutors converge on local conceptual models of an experimental situation. But how do such local convergences become conventional practices in a broader community of speakers? To address this question, Garrod and Doherty (1994) designed another version of the experiment investigating the propagation of conceptual models in a larger communities of speakers. They thus compared maze game dialogues from dyads assigned to two different groups. The first group consisted of “isolated dyads” (like in the previous version of the study) that each played nine trials of the maze game. In the other group, hence the “community group,” each member completed the same number of games as in group one, but on each occasion with a different partner chosen from within the group. In this way, the second group would simulate a simple linguistic community. The purpose of this manipulation was to compare how and to which degree the two groups converged on a common conceptual model of the situation.

The results can be summarized as follows: First, the speakers in the community group rapidly converged on a single model. In the last rounds of games only the matrix model was used (and in fact a specific version of the matrix model where the rows are marked by letters, a, b, c, etc., and the lines are marked by numbers). This is in contrast to the isolated pairs where there was an almost equal distribution of the line and the matrix model. Also the path and figurative model were used in the last games but to a lesser extent. Second, the authors calculated the inter-speaker coordination operationalized as the likelihood that a speaker’s choice of description for one exchange can be predicted from the partner’s choice in the preceding exchange. They found that for the community group this probability is initially low, but increased dramatically during trials ending at almost 1. This is in stark contrast to the isolated pair group. Here the probability was higher in the first games since they were playing with the same partner. But
for the rest of the games no improvement was found, in fact, there was rather a
decrease in co-ordination towards the end where the probability dropped. The
strong coordination found in the community group seems to approach something
like a convention for a particular way of conceptualizing the situation. From this
follows the prediction that participants of the community condition would be less
likely to let local salience influence their description. The test case is a conflict
situation where the local salience affords a construal other than the established
conceptual convention. This is found when, for instance, a participant is ap-
proaching the target and the speaker says something like “the goal is two to the
left from where you are now,” rather than “the goal is at c3.” Garrod and Doherty
found that participants of the community group were in fact less inclined to real-
ize such local affordances and thus divert from the conventionalized conceptual
model.

If we look more globally on the findings of Garrod and Doherty, some interest-
ing questions arise: How can we explain the rapid convergence on one common
model in the community case? If speakers enter a new game with conflicting con-
ceptual models, what then determines on which they will converge? The solution
seems to lie in the local history of preceding games. If, in earlier games, interlocu-
tor A has used the matrix model ten times and the line model only two times,
whereas interlocutor B has used the line model seven times and the matrix model
five times, then on a global level the matrix model has been dominant and will be
more likely to win in subsequent exchanges. Not unlike the “principle of prece-
dence” reported for intra-pair dynamics, this can be formulated as a general prin-
ciple working on a inter-pair level accounting for how linguistic constructions
can propagate in a larger community: whenever there is a conflict between two
models the one that was most commonly used in their immediately preceding verbal
interaction will dominate (Garrod and Doherty 1994: 202). Once stabilized in the
community, a conceptual model come to constitute a common ground and as
such serves as a coordination device for a recurrent problem solving in similar
situations (which conforms to H. H. Clark’s redefinition of the concept of conven-
tion, see Clark 1996).

2.3 Joint assessment of subjective experience

While both the emergent linguistic categories and the conceptual models of refer-
ence considered above are found to rely on motivated realizations of jointly per-
ceived, salient properties of the stimuli space, other kinds of meaning-making
seem – at least at first sight – to challenge such approaches. In the following, we
will thus focus on dialogical processes involved in conceptualizing properties of
perceptual experiences themselves. That is, cases in which we do not attend to objective properties of the environment as much as to the qualities of our subjective experience of those properties. The conceptual relation in these cases is thus not between two or more entities in the environment, but rather between two or more experiences of such entities. Besides, while the categories and conceptual models treated in previous sections tend to be of the discrete kind, these experiential qualities are rather a “matter of degree.”

Many everyday contexts of joint action and coordination depend on our abilities to share such inherently subjective judgments, appraisals, and assessments. In such cases, interlocutors must establish conceptual models to scale and assess individual experiences and performances whereby making it possible to publicly share and integrate their informational content. In the following, we will investigate two aspects of this kind of introspective, scalar meaning: the object of introspection (e.g., assessment and communication of pain), and the degree of confidence in such introspections.

Being able to publicly express introspective assessments is not a trivial task. For interlocutors to successfully communicate introspective assessments or degrees of certainties they must first converge (explicitly or implicitly) on a shared scale for doing so. This is evident from the literature on pain assessment. Communicating about pain to peers and medical personnel is a quite common experience that puts interlocutors in the situation of having to share an otherwise private experience (Lascaratou 2007). It is however difficult to describe the feeling of pain. This is not necessarily because we lack the words: there is an abundance of pain-related verbal expressions, many metaphoric and most of them unusual (Turk and Melzack 2001). But the linguistic expressions people employ in their descriptions of pain may not be effective due to differences in “languages,” experiences, and frames of reference. Indeed, research on linguistic self-reports of pain has deemed them to be highly unreliable since they show unpredictable variations due to gender (Strong et al. 2009), cultural, religious (Waddie 1996), and anamnestic history (Blanchard et al. 1982; Urban et al. 1984), as well as just idiosyncratic factors (Williams et al. 2000).

---

8 Language thus makes possible a conceptual scaffolding that was not in the individual before the interaction, resembling Vygotsky’s notion of internalization: “Every function in the child’s cultural development appears twice: first, between people (inter-psychological) and then inside the child (intra-psychological). This applies equally to voluntary attention, to logical memory, and to the formation of ideas. All the higher functions originate as actual relationships between individuals” (Vygotsky 1978: 57).

9 By confidence we mean “a person’s strength of belief about the accuracy or quality of a prediction, judgment or choice” (Peterson and Pitz 1988; Sniezek 1992).
Given that pain measurement is an important diagnostic resource, a lot of effort has gone into developing and testing standard questionnaire procedures for reporting pain (Anciano 1985). The attempt has been to create cultural artifacts that constrain the linguistic expression of pain in intersubjective situations.\(^{10}\) The construction of such scales of commensurability is at work not only in sharing the contents of introspection, but also in assessing the reliability of the reported information (Peterson and Pitz 1988). Reliability is one of the key factors in modeling cognitive processes (Kruschke 2010) and indeed when confidence ratings accompany information in contexts of decision-making their impact seems to be uncontroversially demonstrated (Pulford 2002; Sniezek and van Swol 1997; Zarnoth and Sniezek 1997) both in adults and children (Moore et al. 1993).

Although less structured than in the case of pain research, attempts have been made in the establishment of a constrained language for confidence assessment of events and states of mind. Given that subjects feel more comfortable communicating probabilities and confidence by using verbal expressions than by using numerical ones (Hamm 1991), studies have been conducted assessing the numerical values attributable to linguistic labels like “absolutely impossible,” “rather likely,” “very probable,” and “almost certain” (for a review and the list of the 178 expressions investigated cf. Clark 1990; Druzdzel 1989). The main conclusion is that although there is a pronounced between-subject variability in the numerical values assigned, the ordinal relations between these expressions are more or less consistent. However, Druzdzel (1989) shows that the numerical and ordinal value of the expressions is also shown to be highly context and task sensitive, with an asymmetry and a large overlap between different phrases. There seem to be several co-existing and partially overlapping uncertainty concepts:

Which answer one will receive to a question about uncertainty or probability will be dependent upon the family of concepts being activated. Thus questions of “uncertainty” may trigger a different line of thought than a question of “chances”; “probabilities” may be assessed independently of questions about “usualness” or “frequency of occurrence”; in hindsight, the appropriateness of “hopes” and “doubts” will be judged by other criteria than their corresponding probability values, etc. (Teigen 1988: 37–38)

---

\(^{10}\) While initially “verbal rating scales” were employed to measure pain intensity, the complexity of pain experience as enacted and expressed by the patients has brought to the development of multi-dimensional questionnaire like the McGill Questionnaire on Pain attempting to better structure and constrain pain descriptions along three dimensions – sensory, affective, and evaluative (Melzack 1975).
It thus seems that the uni-dimensional assumptions in the studies reviewed by Clark and Druzdzel tend to obscure the rich and context-sensitive semantics of expressions of confidence (Fox 1986).

Following the assumptions of this article, such semantics can be better unveiled by analyzing how subjects spontaneously produce and align linguistic ways of expressing and negotiating confidence. Such situations have recently been experimentally assessed by Bahrami et al. (2010). Pairs of participants were separately shown two faint visual displays. The participants had to individually evaluate which of the displays contained an oddball. If their individual choices were divergent, the participants were to negotiate, by freely discussing with each other, a joint decision. Surprisingly, the result was that dyads tended to develop a joint accuracy that was significantly higher than the accuracy of the best individual dyad member. Analyses of the dyads’ linguistic interactions (Fusaroli and Tylén 2012, Fusaroli et al. 2012) show that in order to reach such an effective level of joint accuracy, it is critical that participants come to share a scalar system of confidence descriptors. While linguistic variation between dyads is generally quite pronounced, there are properties of the communication that are highly predictive of the success of the group. Not unlike the findings of Galantucci (2005), Garrod and Anderson (1987), and Garrod and Doherty (1994), within-group convergence on shared lexical repertoires for talking about the perceptual experience and certainty helps establish a conceptual scale (common ground) on which to map confidence intervals. In other words, dyads that settle on a specific (for that dyad) set of graded linguistic descriptors perform significantly better than dyads indecisively drifting between different sets or using individual (non-shared) lexical repertoires.

Common to the cases of pain and confidence assessment is the negotiation or imposition of a linguistically constructed conceptual scale. Such conceptual scales thus impose structure on interlocutors’ introspective experiences and allow them to publicly share experiential qualities.

3 Discussion

A central premise of the framework drafted in this article has been that a theory of “ideal” or “universal” aspects of linguistic meaning should also be able to account for 1) the grounding of these invariant aspects of linguistic meaning, and 2) the pronounced synchronic and diachronic variability found in language. Classic positions founded on the works of, e.g., Frege – and to some extent Saussure – do not satisfy such premises. While embodiment theory, as formulated, e.g., by
Lakoff and Johnson, has made promising progress on these issues, it still seems ignorant to the central constitutive role of social semiotic dynamics (Baronchelli et al. 2005; Galantucci and Garrod 2010; Thibault 2004), and seemingly has no solution to issues of language change.

Central to the dynamic approach we have pursued here, is an emphasis on the situated socially interactive component of meaning-making. It is argued that invariant patterns of meaning are motivated and stabilized through an interplay between structural stabilities bound to our biological bodies, physical and cultural environments, and local dialogical processes in which interlocutors pursue situated communicative needs. Our physical, biological and cultural environments thus constitute a rich and multifaceted set of structural affordances and constraints that guide and scaffold online human meaning-making. These structural affordances constitute the universal, invariant ground for linguistic meaning. But it is only in concrete situations of contextualized language use that certain structures are selected for and realized in linguistic categories, construals, and conceptual models. The selective realization of some structural affordances rather than others gives socio-cultural variability to human meaning-making. Synchronic variability can be motivated by local environmental saliences (e.g., the morphology of the maze, as in Garrod and Doherty 1994, or the colour spectrum encountered in a special environment as in Puglisi et al. 2008), but likewise by the local history of social interactions. Interlocutors thus largely orient to their experiences of recent successful communications and these experiences thus come to constitute another set of affordances working on equal footing with physical, biological, and cultural structural potentialities. In fact, as patterns of linguistic interaction spread and stabilize in a speech community, they tend to conventionalize. As a consequence, certain linguistic practices can overrule otherwise strong local affordances of referent scenes (as discussed in Garrod and Doherty 1994) due to their normative inscription in the community. Such mechanisms thus seem to ground diachronic and typological variations among the world’s languages.

References


**Bionotes**

**Kristian Tylén** (b. 1974) is a postdoctoral researcher at Aarhus University (Kristian@cfin.dk). His research interests include dynamical approaches to language and cognition, cognitive aesthetics, and experimental semiotics. His publications include “Language as a tool for interacting minds” (with E. Weed et al., 2010); and “Coming to terms: Quantifying the benefits of linguistic coordination” (with R. Fusaroli et al., 2012).

**Riccardo Fusaroli** (b. 1981) is a postdoctoral researcher at Aarhus University (fusaroli@gmail.com). His research interests include linguistic coordination, experimental semiotics/pragmatics, and complex systems approaches to social
interactions. His publications include “Coming to terms” (with B. Bahrami et al. 2012); “The self-organization of human interaction” (with R. Dale et al., in press); and “Dialogue as interpersonal synergy” (with J. Raczaszek-Leonardi & K. Tylén, in press).

Peer F. Bundgaard (b. 1967) is an associate professor at Aarhus University (sempb@hum.au.dk). His research interests include aesthetic semiotics and the crossovers between cognitive linguistics and phenomenology of language. His publications include “Towards a cognitive semiotics of the visual artwork” (2009); “Husserl and language” (2010); and “The grammar of aesthetic intuition – on Ernst Cassirer’s concept of symbolic forms in the visual arts” (2011).

Svend Ostergaard (b. 1949) is an associate professor at the University of Aarhus (semsvend@hum.au.dk). His research interests include cognitive linguistics and cognitive neuroscience. His publications include The mathematics of meaning (1997).