

• subclausal: prospective or retrospective increments, semantically belonging to a clause but prosodically isolated into a separate EDU; see example (2).

- (2) yats'ese di'isdiyok dine k'inodle ghoda
that's.why it.happened.to.me that time icon because.of
'That is why that happened to me then because of the icon'

Long EDUs primarily consist of combinations of a matrix clause and a complement:

- (3) hondenh ghwla' sidadza' yinezinh ts'e'
where unknown my.sister he.is.thinking Particle
'Where is my sister, he was thinking'

Much rarer are concatenations of coordinate clauses within one EDU, or relative clause constructions.

Generally, the stratification of EDUs in UKA is quite typical, judging by the data we have from better studied languages. Probably the most surprising fact is the equifrequency of short and long EDUs. For comparison, in the Russian corpus studied in Kibrik and Podlesskaya 2009, short EDUs strongly outnumber long EDUs: 26% vs. 6.3%. Most likely, this peculiarity of UKA is related to its polysynthetic character. If measured in the number of words, EDUs in a polysynthetic language are shorter: more information is packed in the inflected verb. As a result, more additional lexical elements fit inside an EDU. There are fewer regulatory and subclausal elements finding themselves outside an EDU, and more than one verb more often fits inside an EDU.

The profile of a language in the domain of local discourse structure thus depends on two major

factors: first, the universal, cognitively based requirements on discourse segmentation, and second, language-specific grammatical peculiarities of the language.

This study is supported by grant #11-04-00153 from the Russian Foundation for the Humanities.

Chafe W. 1994. *Discourse, consciousness, and time*. Chicago: University of Chicago Press.

Collins R., Petruska B. 1979. *Dinak'i (Our words)*. Upper Kuskokwim Athabaskan Junior Dictionary. Anchorage: NBMDC.

Iwasaki S., Tao H.-Y. 1993. *A comparative study of the structure of the intonation unit in English, Japanese, and Mandarin Chinese*. Paper presented at the annual meeting of LSA.

Kibrik A. A., Podlesskaya V.I. (eds.) 2009. *Rasskazy o snovidenijax: Korpusnoe issledovanie ustnogo russkogo diskursa* [Night Dream Stories: A corpus study of spoken Russian discourse]. Moscow: JaSK.

Matsumoto K. 2000. Japanese intonation units and syntactic structure. *Studies in Language* 24: 525-564.

Wouk F. 2008. The syntax of intonation units in Sasak. *Studies in Language* 32: 137-162.

NON-DISCRETE EFFECTS IN LANGUAGE, OR THE CRITIQUE OF PURE REASON 2

Andrej A. Kibrik

aakibrik@gmail.com

Institute of Linguistics, Russian Academy of Sciences (Moscow, Russia)

Language is a hierarchical system. At each hierarchical level (phonology, grammar, discourse) units display a paradoxical behavior. They are segmental, and at the same time they somehow tend to avoid segmentation and merge. This can be seen in both paradigmatic and syntagmatic aspects. Consider the phonemic level. Paradigmatically, each language is typically believed to have a fixed set of phonemes. But all kinds of partial membership in this set are systematically found across languages, e.g. it is unclear whether one

must posit the difference between hard and soft /k/ and /kʲ/ in Russian. Syntagmatically, it is difficult to draw a clear boundary between segments in phonetic signal. For example, when pronouncing something like /ko/ labialization is found already when the consonant is pronounced.

In grammar, the neat distinction between words and affixes is hard to be drawn in any language. Elements such as English *the* or *to* are words by some criteria, and parts of larger words by other criteria. Linguists typically attempt to solve this problem by introducing an intermediate class of elements: "clitics". But this actually complicates the problem even more, as now one has to draw two boundaries: between words and clitics and between clitics and affixes. At the level of discourse structure,

there is evidence that language is produced not as a steady flow but rather as a sequence of spurts, or “elementary discourse units”. These units can often be identified by prosodic criteria, but there is always a residue of complicated instances in which it remains unclear whether we see a combination of elementary discourse units merging together or, on the contrary, an unusually long but single unit.

There are multiple non-discrete effects in semantics, too. The well-known phenomenon of polysemy, found in most lexical items, resists discrete analysis as boundaries between different meanings of a word are blurred, and seemingly distinct meanings of a word can sometimes be realized simultaneously (see Zaliznjak 2006 for a detailed discussion).

Overall, it appears that language simultaneously longs for discrete, segmented structure and tries to avoid it. This problem is truly overwhelming in linguistics and is in the core of theoretical debates about language. Non-discrete effects permeate every single aspect of language. Two polar approaches to this problem are found. The dominant approach can be called discrete or digital, it suggests that language is underlyingly discrete whereas non-discrete effects only occur at the stage of realization and therefore are uninteresting. The digital approach is reductionist but it traditionally has an appeal of apparent scientific rigor.

The other approach, that can be dubbed analog, acknowledges non-discrete effects in language and, in trying to grasp them, has developed theoretical concepts such as prototypes in category membership. The analog approach is inclusive but repels some by picturing reality in an overly complex way. Among the many attempts in linguistics and in cognitive science to capture non-discrete effects consider e.g. Rosch 1973, Bell 1976, Lakoff and Núñez 2000, Linell 2005, Nikolaeva 2008. Varieties of the analog approach are very diverse and often keep struggling with the traditional opposition “discrete vs. continuous structure”. I propose that in the case of language we see the third kind of structure that can be called focal: focal phenomena (phonemes, meanings, etc.) are simultaneously distinct and related.

All this poses a question about the scientist’s inclination to see language as a discrete structure: is this inclination based on the objective properties of language or rather is a product of the observing human mind? This kind of question is not novel in the history of scientific thought. In fact, it is one of the main issues addressed in Kant’s *The Critique of Pure Reason*. Kant attempted to go beyond the debate between empiricists and dogmatists and suggested that the position of observer, or

cognizer, fundamentally affects the knowledge of the world. If language is a *Ding an sich*, according to Kant it must be unknowable. Only those things are knowable that are open to experience and correspond to our intuitive judgement. Of course, compared to physical world, in the case of language and other cognitive processes Kant’s problem is much more acute because mind here functions both as observer and object of observation, so making the distinction between the two may actually be impossible.

Standards of scientific thought have developed on the basis of physical, rather than cognitive, reality. Of course, physical reality is much more prone to the discrete approach. Just as we invoke scientific thinking, we tend to immediately turn to discrete analysis, and this is the reason why discrete linguistics is so popular, in spite of the omnipresence and obviousness of non-discrete effects. Much of science is about categorization of phenomena, and it is an extension of the general cognitive ability to categorize that is often associated with the left hemisphere functions, rationality, and discreteness. However, in the case of language and other cognitive processes we clearly see the limits of the traditional discrete approach. We need to develop a more embracing linguistics and cognitive science that address non-discrete phenomena not as exceptions or periphery of language and cognition but rather as their core.

I propose several avenues of research leading towards this goal. First, it makes sense to begin developing linguistic analysis from the kind of sound channel that is far less discrete than the verbal code, namely prosody. One example I consider is connected to the identification of the “period intonation” in speech (Kibrik 2008). This prosodic pattern means the movement of the speaker’s F0 towards the bottom of the speaker’s voice range. This definition inherently relies on a speaker’s characteristics, it cannot be formulated in more objective terms, such as a certain F0 value. Other prosodic examples are also considered, such as vowel length iconically representing physical or cognitive distance (Kodzasov 2009).

Second, another communication channel that is also essentially analog, is gesticulation accompanying speech. As was rightly pointed out by Tomasello (2008), in order to “understand how humans communicate with one another using a language <...> we must first understand how humans communicate with one another using natural gestures”. Gestures are typically classified into several kinds, including pointing gestures and iconic gestures. There are many instances in which one and the same gesture is simultaneously pointing

and iconic, e.g. when a speaker demonstrates the motion of vehicles, using a hand as a substitute of a vehicle and at the same time pointing in a certain direction with his or her fingers.

The third point is methodological. A lot of efforts were spent beginning from the 1960s trying to explain language mathematically as a discrete symbolic system. These efforts have largely failed, mostly because of the non-discreteness of many linguistic phenomena. It is time to think of alternative kinds of mathematics more appropriate to the nature of language. I describe a study of referential choice in discourse, in which the choice between full and reduced noun phrases is seen as not necessarily categorical, and methods of machine learning are used to model this process and assess probabilities of a certain referential option appearing in discourse. This last point demonstrates that not only non-verbal (prosodic and gestural) but also verbal devices can be explored, while acknowledging non-discrete effects in language.

It remains an open question if cognitive scientists are able to eventually overcome the strong bias towards traditional rationalism and discrete analysis, characteristic of scientific thought. This bias may be a cultural tradition or it may be

something deeply grounded in the human analytical mind. Anyway, it is worth trying to circumvent this bias and to seriously explore the focal, non-discrete structure that is in the very core of language and cognition.

This study is supported by grant #11-04-00153 from the Russian Foundation for the Humanities.

Bell R. 1976. *Sociolinguistics: Goals, approaches and problems*. London: Batsford.

Kibrik A.A. 2008. Est' li predloženie v ustnoj reči? [Is there sentence in speech?] In A. V. Arkhipov et al. (eds.) *Fonetika i nefonetika. K 70-letiju Sandro V. Kodzasova*. Moscow: Jazyki slavjanskoj kul'tury, 104–115.

Kodzasov S.V. 2009. *Issledovanija v oblasti russkoj prosodii* [Studies in Russian prosody]. Moscow: JaSK.

Lakoff G., Núñez R. 2000. *Where mathematics comes from*. New York: Basic books.

Linell P. 2005. *The written language bias in linguistics*. London: Routledge.

Nikolaeva T.M. 2008. *Neparadigmatičeskaja lingvistika* [Non-paradigmatic linguistics]. Moscow: JaSK.

Rosch E. 1973. Natural categories. *Cognitive Psychology* 4: 328–50.

Tomasello M. 2008. *Origins of human communication*. Cambridge, MA: MIT Press.

Zaliznjak A.A. 2006. *Mnogoznačnosť v jazyke i sposoby ee predstavlenija* [Polysemy in language and ways to represent it]. Moscow: JaSK.

ON THE PRAGMATIC AND BIOLOGICAL NATURE OF LANGUAGE DYNAMICS

Pablo Kirtchuk

kirtchuk@vjf.cnrs.fr

LACITO-CNRS, INaLCO (Paris, France)

This is part of LUIT (*Language, a Unified and Integrative Theory*, forthcoming, cf. Kirtchuk 2007). The dynamics of language involves diachrony but also, among others, ontogeny, phylogeny, creolistics and register variation; and not only grammar but first and foremost pragmatics. For instance: Pragmatic, intonative, morpho-syntactic, typological and psychological factors, show that more often than not, so-called proleptic utterances do not result from the extra-position of elements from sentences previously constructed. Terms such as *extra-position* and *dislocation* imply the precedence of syntax over pragmatics and over language's nature, which is multidimensional and cognitive and not merely grammatical. Language is not dynamic only as a phenomenon, even its actual manifestations *function* dynamically and each one of them reflects the properties of language as a whole. In this sense, language is a **fractal**. Even terms as 'grammar or structure of information' are misleading inasmuch

as they imply a structure, while the *raison d'être* of so-called proleptic and 'dislocated' utterances is reflecting a natural iconic pragmatic order relatively independent of the constraints imposed by the structure of the language in which those utterances are produced. Proleptics are narrowly akin to topic-first utterances, which are spontaneous and as such require a minimal encoding and decoding effort, while grammatically well-formed sentences must conform to grammatical rules, especially of word-order and agreement. Proleptics often include the presence of a co-referent element both in the main and in the subordinate clause, most often with some kind of agreement, so they include a morpho-syntactic component, while topic first utterances most often do not. Yet there is an affinity between the dynamic parameters of utterances with focalization or topicalization and of so-called proleptic ones, too consistent to be imputed to coincidence alone. As they are founded on pragmatic and communicative factors, proleptics precede their syntactically well-formed, *i. e.* grammatical vis-à-vis, of which they are the second stage in the gradual displacement from the pragmatic to the grammatical mode.