

## Conclusion

Once fact that emerged from the analysis of interactions is that social communication is a matter of dynamic coordination of multimodal units of semiotic meaning in languaging. Rigid positions like the ‘autonomy of syntax’ (Chomsky, 1965, 1972) should find ways to accommodate new conceptions of language, in which grammar is ‘emergent’ Hopper in (Tomasello, 1998: 155-175); and “brings in local flexibility to revise and adjust gestalts while they are being put into speech” (Auer, 1996: 75). Repair is inbuilt in the very structure of talk in interaction (Schegloff, 1987); in the very dynamic nature of the grammar of sentences in progress (Lerner, 1991). In this perspective, grammar and “prosody, cooperate in very delicate ways, each of them on the basis of their semiotic possibilities” (Auer, 1996: 75; cf. also Couper-Khulen and Selting, 1999, 2006), so that “the meanings contributed by lexical and grammatical elements evoke [and constrain] – but do not fully constitute – the overall conceptualization evoked by a complex expression”. [The composite unit] is an entity in its own right with many extra sources of compositionality” (Langacker, 2009: 4, in Tomasello 2009, emphasis added in parenthesis).

More crucially, these perspectives on ‘language’ subsume that all constituencies interact in a ‘conceptually blend mental space’ (Fauccounier and Turner, 1994; 1996, 2008; Turner and Fauccounier, 1995), whereas ‘lexicon, morphology and syntax prosody form a continuum fully describable as assemblies of symbolic structures” (Langacker, 1987a, 1990, 1991; cf. Langacker in Tomasello, 1998: 2). That the constituencies lexis and syntax interact ‘in a blend space’ is made clear from examples of lexical omissions whereby a full lexical null influences the syntax of the overall sentence. However, not only constituencies of language (Halliday, 2004) but also multi-dimensional multi-modal) variables such as pitch and facial movements interact in the same packaging. Languaging itself is composite multimodality and speakers produce ‘multimodal utterances’ (Hutchins and Nomura, 2001). Developmentally speaking (cf. Eibl-Eibefeldt, 1972; 1973 and Tomasello, 1999), there exist a point from which everything grows and (McNeill, 1995, 2005). This point is the specifically enchoric (Enfield, 2009, 2011) locus of ontogenic existence in a situated cognizer’s Innenwelt (von Uexküll, 1957) that is subsequently developed in composite social semiotics (Halliday, 1978; Hodge and Kress, 1988; Lemke, 1997; Kress and van Leeuwen, 2002) on the social plane of reality (Rogoff, 1991), through talk-in-interaction (Sacks et al., 1974).

Accessing this locus of ontogenic existence is – luckily – practically impossible. Thus languaging and language are not at all predictable matters. As a matter of fact, an optimal or a universal combination for the contextualisation of repair does not exist. There exists a tendency for action units and sometimes pitch movements to correlate in certain instances of certain repair types, for example we found maximal pitch movements and AU 1+2 or alternatively AU 1+4 and AU 1+2+4 combinations on correction proper and other – increment – repair proper like replacements to an host. This goes in line with previous research on facial expressions and emotion (Friesen and Ekman, 1983) according to which movements of the frontalis medialis and lateralis occur – it seems – universally in the emotions of fear and surprise. As a matter of fact, a speaker/cognizer (cf. Tomasello, 1999) may rightfully be both surprised and alerted to have committed an error. Physiological reactions associated with facial expressions have previously been discussed by Ekman, Levenson and Friesen (1983) who for example found greater heart acceleration rate coincided with facial expressions underpinning strong emotions such as fear. Signs are conventionally indexical. This however does not imply that signs produced are ‘right’ or true, nor that they are easy to interpret and discern.

As far as the analysis here conducted, the authors reckon findings all but conclusive. This is the case because there are several imprecisions and bias in the methodology of this work, for example some theoretical and pragmatic variables that were neglected in the process but ultimately found to be very influential in the contextualisation of repair. Some of those encountered that make any theoretical claim in this work (and perhaps some related work too) de facto uncertain are: a) some facial micro-expressions are articulatory, thus relate to the anatomy of speech and the physiology of languaging; and especially in the case of current speakers in talk in interaction, it is often impossible—even with physiological data – to tell an articulatory from a non-articulatory expression; b) some facial micro-expressions are triggered by gaze and other body behaviour, for example the muscles frontalis lateralis and medialis are contracted upon blinking the eyes or on head thrusts, thus further hindering the exact classification in terms of ‘voluntary and involuntary’; c) age and loss of muscular tone may have an impact on the quality, intensity and duration of micro-expressions thus destroying the solidity of any empirical claim in the neurophysiology of facial semiotics. These and more are expressions of diversity that because existing in theory, must be accommodated in praxis. To conclude on this, the author also wishes to state that the fact these variables were neglected before

conducting the study is exclusively the authors' fault that has evidently run an approximate pilot analysis of the phenomenon, eventually resulting in the development of a biased methodology.

The work is also biased in annotations and transcription as it did not (and quite a few works in social sciences do not) take in consideration the margin of error. It has happened to me that the more micro-analytically I zoomed in transcripts and phenomena, the bigger margin of error and diversity I found. I also found that it does not really matter how many cycles of segmentations, transcription and annotations one performs, because the margin of error cannot be eradicated. Never. Eventually, this has a huge impact on the interpretation of data, and makes any theoretical claim uncertain, fallible and intrinsically biased.

After years of researchers' dissertations on the topic, the issue of why speakers commit errors is to my best understanding still shrouded in mystery. The fact that speakers cut off upon detection of errors, and in some cases even before uttering them, shows not only that cognizers undergo momentary dissociation in their cognition, but also that speakers are capable of pre-articulatory monitoring to re-solve dissociation in 'correction' eventually resulting in the formulation of their intended plan (Butterworth, 1992). However, because the locus of ontogenic existence of a semiotic composite remains in the private property of the cognizer, it is –to my best understanding – not yet clear when monitoring starts and how the process unfolds. So, while previous research (Allen et al., 2007; Ozurek et al., 2008) analysed cross-linguistic variation in semantic coordination of speech and gesture and found evidence for an interface representation of spatial thinking and speaking; issues with time (and monitoring) are still unresolved. As a matter of fact in the field of social science, and more specifically in the domain of speech with gestures composite, researchers admit that many inconsistencies exist and present contrastive results as for the synchronicity of speech and gestures composites. Similarly, in research on the duration of facial expression, with duration being the main feature that distinguishes facial expressions from micro-facial expressions, it is not yet clear how long a micro-expression is and with such an uncertainty, it is not possible to talk about voluntary and involuntary aspects of cognition, and very difficult to formulate any claim based on the analysis of these notions (and perhaps on other notions too).

Finally, it has been brought to my attention that it is perhaps worth discussing the results of this study in the light of Skinner's (1938) stimulus-response theory, because a bunch of

facial expressions are below 500 ms, thus supposedly involuntarily leaked; outside of a speaker's control. In Skinner's (1938) theory, a stimulus is coupled with a response in one such way that the higher the intensity and the longer the duration of a stimulus, the higher the intensity and the longer the duration of its coupled response too. In the case of the analysis run in this work, stimulus-response theory does not apply for at least three reasons. First is the issue with the duration of the stimulus. While it is often possible to time the micro-facial expression that contextualises repair in response to a speech error so (what here I call 'outcome' with approximation because we never consider the margin of error in calculations), it is not possible instead to calculate the duration of a stimulus (i.e. error), thus impossible to say how far in planning, from the growth point of ontogenic existence to the phylogenetic development of the multimodal utterance the speaker has gone upon detecting an error. This is suggested by the fact that there are units that are cut off for seemingly no reason. We may assume these units are cut off due to pre-articulatory detection of an error, or to accommodate new stimuli. Second and more crucially, there are issues in the detection of a growth point as stimulus itself because a growth point may develop in response to another environmentally coupled stimulus (or probably more stimuli) which we wouldn't be able to fully disambiguate. Third and most importantly, because untraceable, the growth point is not a unit that can be empirically analysed.

From a social perspective, I seem to understand that the fact society develops signs does not necessarily mean these are true or correct in any strict or absolute logic sense. On the other hand however, signs are indexical conventions which are essential for society and earthlings to communicate thus exist.

From the perspective of social psychology, it is perhaps re-assuring to know that, in interactional grammar(s), repair mechanisms are universal, and the chance to correct errors is available, in theory, to everyone, however, in praxis, not to everyone. As a matter of fact, while it tends to happen that speakers decide to repair their own errors, speakers are also found to leave them uncorrected or have addressees and recipients correct them. If, when, who and how errors will be corrected is another unpredictable matter of – individual and social – cognition. Repair is extremely powerful resource speakers may also use to deceive, mask, and conceal information, which is tendentially bad but nevertheless often a moral obligation in the ethical frames of social inter-action.