

Joint Speech as an Object of Empirical Inquiry

Fred Cummins
UCD School of Computer Science
University College Dublin
fred.cummins@ucd.ie

October 24, 2018

This is a pre-print of the article published in the journal *Material Religion*. The special issue can be read online at <https://www.tandfonline.com/eprint/qVSSVjSetT5dkV6UvFwa/full>

1 The Challenges of Joint Speech

Any time multiple people utter the same words at the same time, we have an instance of Joint Speech. With this simple definition, we bring into being a rather odd and challenging object of empirical study. There is no shortage of readily available primary data that satisfies the criterion. We find examples in churches and temples, of course, but also in the secular auditoriums where oaths are sworn and allegiance is pledged. The definition also picks out the repetitive chanting of protest in which collective aspirations and passions are made manifest. To these, we have to add those tribal practices in which group identities are enacted among sports supporters (some sports more than others), and a motley variety of educational and informal practices including such trivial rituals as the singing of Happy Birthday. There is plenty of joint speech to study, and participation in such activities in some form or another appears to be ubiquitous through millennia and across the globe.

But any approach to studying joint speech as an important form of human activity spanning all these domains runs up against a poverty of frameworks, concepts or theories that might be brought to bear (Cummins, 2014a). Many of the distinctions we implicitly rely upon as we approach language fail when applied to joint speech. There is no fundamental distinction between speaker and listener, so any discussion of this language behavior as communicative message passing is already in difficulty. The texts used are authored elsewhere—something that Rappaport (1999) has noted is central to rituals, but that is starkly different from the creativity and generativity of language as conventionally portrayed. There is no neat divide between speech and music. Repetition is a very common feature of joint speech, and together with the production of synchronized gestures, this leads to rhythmic enhancement and melodic stylization of the repeated texts. The English word “chant” captures well this ambiguity between speech and song.

Perhaps most saliently, joint speech does not belong on the page. Many of the characteristics of language that have received the attention of scientists are precisely those that survive the translation from the voice to writing, including such matters as syntax (grammar) and lexis (choice of words). But joint speech is not suited to this kind of analysis. Rather, participation in a situation of co-presence seems to be absolutely central to the purposes which joint speech serves (Cummins, 2018),

however we might understand those to be.

In joint speech we are faced with the vocal manifestation of collective aspirations, collective purposes, and collective sentiments. There is no way to understand joint speech that does not address this head on, and therein lies an immense challenge to scientific practice. For intentionality has by and large been treated as a matter for individual minds belonging to solitary individuals. An uneasy equilibrium exists in the division of explanatory competence between the psychological sciences (resolutely and fundamentally individual) and the social sciences (collective, but lacking ontologically secure foundations). Since the structuralist and cognitivist characterisations of language as an abstract self-contained system (*langue*, syntax), the intentional content of utterances has been consigned to the hidden domain of the individual mind, or cognitive system. There is not yet any secure scientific framework for characterizing a form of language that resists reduction to the hidden minds of individual persons, and that demands acknowledgement of the intentional being of a collective.

2 Using Joint Speech to Direct Attention

With these observations, an opportunity arises. The simplicity of the empirical definition serves to establish *prima facie* links between these several domains. We may lack a definition of “ritual” that commands general assent, but by using joint speech as a means of focussing our attention, we might evade some definitional issues and thereby motivate the joint consideration of observations from church, stadium, classroom, and the protest march. Let us list some features we find that transcend any one of these domains.

Repetition is such a frequent characteristic of joint speech, that its absence, when it occurs, is itself noteworthy. In practices of prayer, mala or beads are used as a material support for counting repetitions, but repetition is found in protest marches and football stadia too. Clearly it is not enough that some words be spoken. Something happens *in the speaking*, providing our first indication that a vocabulary of *enaction* rather than production might be appropriate here.

We find a tight integration of gesture and speech. Although the codification and choreography differs, we find a strong association between repeated chants and associated gestures of hand, legs and trunk, in all of the major domains in which joint speech occurs. In highly preserved ritual and rite, such gestures frequently become explicitly choreographed, stylised and saturated with significance. Joint speech practices encourage us to view such speaking as a strongly *embodied* activity.

Participation is central. Whether swearing an oath, calling for the downfall of the regime, or reciting a *credo*, the act of uttering together with others is central. A limiting case might be adduced, which is the minute’s silence observed at public gatherings in the aftermath of tragedy. Here, the words have been reduced to an absolute minimum, yet the collective activity is sustained by the simple expedient of participation.

Where joint speech occurs, there is frequently a complex division of labor between a leader and the larger crowd. Crowds have a reputation for lacking the finer discriminatory capacities of reason. Yet we find that complex texts expressing nuanced points may be employed, particularly in ritual and in political gatherings. In each case, the task of speaking the text that must be enunciated carefully, whether it be for liturgical or political reasons, falls to the leader, but the assembled crowd gets to partake and thus participate in the expression through the use of a generalized expressions of assent: “amen” in a religious setting, “right on!” or “yeah!” at a political rally. Viewed in this

light, the carefully curated structures of liturgy may be considered as continuous with forms of organization found in secular and improvised contexts, thereby affording an *-etic*, rather than an *-emic*, perspective.

Thus joint speech may make us aware of more-or-less generic features of collective intentionality that transcend domains, and may thereby illuminate and enrich our understanding of how such features become elaborated within specific domains. As the microscope serves the study of microbes, so joint speech may serve the study of ritual and the founding of collectives, not by reduction or explanation, but by directing our attention.

3 Joint Speech and Enaction

With its themes of strong embodiment, the performative enactment of identity, and the arising of normative perspectives through synchronised action, the vocabulary of the theory of enaction may be of use in further exploring joint speech and associated activities (Stewart et al., 2011; Cummins, 2013). This is not yet a mature scientific framework, however, and there is considerable variety in the way the theoretical vocabulary is applied. To some, it provides an alternative way to pose questions that are answered, with rather different assumptions, within an individualistic cognitive psychology. Those theoretical approaches that emphasize sensorimotor engagement with the environment are relevant here; the “organism” or “system” that features in such accounts is usually the single body, considered as co-extensive with (if not reducible to) the person, and the explanatory targets include accounts of skilful action and of perception (O’Regan and Noë, 2001). But enactive theory has long aspired to more generality than this. Some of its roots lie in a foundational theory of biology, autopoiesis, that considers the smallest living unit, the cell, as its paradigmatic example (Maturana and Varela, 1987; Cummins and De Jesus, 2016). Here, the focus is on the smallest set of mutually related processes that can be said to embody a normative perspective on the world. In the regulated exchange of matter with its surround, the cell enacts its being, and all such exchanges are meaningful from the perspective of the cell, in that they might be said to be supportive of, or detrimental to, that fundamental project of continued being.

Further elaboration of this basic theoretical vocabulary has sought to provide an account of meaningful being-in-a-world that spans all levels at which the processes of life are manifest, from the cell, through individual organisms and their groupings and societies, to the whole of the biosphere (Froese and Di Paolo, 2011; Cummins, 2018). This flexibility with respect to the notion of a normative perspective, which can be characterised in similar fashion for cells, bodies, flocks, tribes and nation states, offers a potential starting point for the treatment of the collective intentionality expressed through joint speech practices.

The self-sustaining activities of diverse kinds of complex systems allow such systems to give rise to concerned perspectives on their worlds, thereby enacting or bringing into being both a generalized “subject” and a concomitant “world,” where this world is to be understood as saturated with significance for the system. This kind of self-producing and self-sustaining activity is generically termed *sensemaking*, and, unlike comparable psychological predicates, the basic idea extends similarly to the activities of both individuals and collectives. When applied to the sense-making activities of multiple individuals in tight reciprocal coordination, we may then speak of *participatory sensemaking* and the enactment of a shared world (De Jaegher and Di Paolo, 2007; Cummins, 2014b). This intricate intertwining of concepts from fundamental biology together with self-sustaining activities that ground collective worldviews refuses any simplistic division between

processes of nature and of culture. But is also demands a kind of hermeneutical self-awareness on the part of scientists doing the describing that is, as yet, largely foreign to the natural sciences, while it may be considered wholly unexceptional in the anthropological and social sciences.

References

- Cummins, F. (2013). Towards an enactive account of action: Speaking and joint speaking as exemplary domains. *Adaptive Behavior*, 13(3):178–186.
- Cummins, F. (2014a). The remarkable unremarkableness of joint speech. In *Proceedings of the 10th International Seminar on Speech Production*, pages 73–77, Cologne, DE.
- Cummins, F. (2014b). Voice, (inter-)subjectivity, and real time recurrent interaction. *Frontiers in Psychology*, 5(760).
- Cummins, F. (2018). *The Ground From Which We Speak: Joint Speech and the Collective Subject*. Cambridge Scholars.
- Cummins, F. and De Jesus, P. (2016). The loneliness of the enactive cell: Towards a bio-enactive framework. *Adaptive Behavior*, 24:149–159.
- De Jaegher, H. and Di Paolo, E. (2007). Participatory sense-making. *Phenomenology and the Cognitive Sciences*, 6(4):485–507.
- Froese, T. and Di Paolo, E. A. (2011). The enactive approach: Theoretical sketches from cell to society. *Pragmatics & Cognition*, 19(1):1–36.
- Maturana, H. R. and Varela, F. J. (1987). *The Tree of Knowledge: The Biological Roots of Human Understanding*. New Science Library/Shambhala Publications.
- O’Regan, J. K. and Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24(5):939–973.
- Rappaport, R. A. (1999). *Ritual and Religion in the Making of Humanity*, volume 110 of *Cambridge Studies in Social and Cultural Anthropology*. Cambridge University Press.
- Stewart, J., Gapenne, O., and Di Paolo, E. (2011). *Enaction: Toward a New Paradigm for Cognitive Science*. MIT Press.