

# The Ground from Which We Speak



# The Ground from Which We Speak:

*Joint Speech  
and the Collective Subject*

By

Fred Cummins

Cambridge  
Scholars  
Publishing



The Ground from Which We Speak:  
Joint Speech and the Collective Subject

By Fred Cummins

This book first published 2018

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data  
A catalogue record for this book is available from the British Library

Copyright © 2018 by Fred Cummins

All rights for this book reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the copyright owner.

ISBN (10): 1-5275-1600-8

ISBN (13): 978-1-5275-1600-7

# CONTENTS

Prologue..... vii

## **I. Introducing Joint Speech**

Chapter One..... 3  
Some Initial Observations

Chapter Two..... 21  
Amen! The Structures of Joint Speech

Chapter Three ..... 35  
Science(s) and Subjects

## **II. The Science of Joint Speech**

Chapter Four..... 57  
The Sounds of Joint Speech

Chapter Five ..... 75  
Joint Speech and Movement

Chapter Six..... 103  
Joint Speech and the Brain

Chapter Seven..... 121  
Language and the Voice

## **III. Beyond Science as Normal**

Chapter Eight..... 151  
An Enactive Starting Point

Chapter Nine..... 175  
Making Use of Joint Speech Observation

References ..... 193

Index ..... 199

## PROLOGUE

The singing of Happy Birthday. The recitation of the Nicene Creed. The countdown to midnight on New Year's Eve. "U.S.A.! U.S.A.!" spontaneously chanted in a high tech control room. The outrage of Syrian citizens during the popular uprising of 2011. The frenzy of "Lock her up!" as chanted during an election campaign. These diverse snapshots open a window onto the use of the voice in making collective purposes and collective identities manifest. They all take place as part of activities apportioned a great deal of significance to those who take part in them. They all make use of the voice in a specific manner: in each case, many people say or sing identical words at the same time. I call this kind of verbal activity Joint Speech.

This book seeks to introduce joint speech as an object of empirical study. In so doing, it uses the empirical study of joint speech to critically examine many assumptions underlying scientific work in those disciplines that deal with the living: biology, psychology, and the social sciences. In the first part the topic is introduced with several concrete examples. The principal characteristics of joint speech are discussed, and readers will be entirely familiar with many of these, as joint speech is an activity all language users partake in. Having established that there is a lot to examine and discuss, a big question then arises: Why has there been virtually no empirical study of this kind of behavior in the human sciences? The absence of a body of scientific work is very revealing, and it points to something of a blind spot. In pursuing this larger question, it is argued that there is an unresolved tension in play about how science should treat subjects, especially collective subjects, when it aspires to some, often unexamined, goal of objectivity. With this, large issues are clearly at stake. Chapter Three considers the way in which subjects and objects become entangled in the sciences of the living, and how joint speech may direct our attention to just those processes in which many of our collective identities are forged.

The middle part of the book then goes on to demonstrate that scientific inquiry of joint speech is both practical and profitable. Worked examples are provided from the diverse domains of phonetics (the sounds of

speech), from movement science (joint speech as synchronized action) and from cognitive neuroscience (where joint speech has some surprises in store). A special consideration is given to how we might think of joint speech within the study of human language more generally. In each case, the scientific work throws up questions about how to handle the divide between subjects of various kinds and their worlds.

Having established that there is a significant absence of scientific work, and then demonstrated that such work is possible and produces novel insights in many different ways, the final section turns back to the difficulties raised at the start. It argues that the study of joint speech might open a window to the empirical study of practices that ground human experience and identity. This may give us a useful and powerful way to approach the study of many kinds of important human activity, and the multiple overlapping collective identities that are thereby brought into being. This points towards a radical reconsideration of what scientific activity is, and how far its truths stretch. To support this ambitious venture, some suggestions are made about how one might appropriately develop a technical language suited to consideration of multiple perspectives, and how one might appropriately handle the relationship between subjects and their worlds.

The empirical phenomenon being discussed, joint speech, should give rich food for thought. As familiar as it is from every day life, whether one indulges in religious rituals, takes part in political protest, or merely chants merrily on the football terraces, it is relatively easy to show that joint speech is a very special kind of language use. It appears to be far older than writing and to have played a role, largely unexamined, in the foundation of all human societies. One might almost question whether it should be regarded as language, in a strict sense, or not, as many things we have come to expect of language, such as the sharp distinction between speakers and listeners, do not seem to apply to joint speech. Indeed, the study of joint speech necessarily leads us to a larger view of just what language is, and how the vocal activity of humans works and how it matters.

The questions raised for science itself may not be as familiar to some readers. In the course of the book, we will have to recognize some commitments within the life and human sciences that are frequently unacknowledged. It will be argued that objectivity in science is a complex issue, especially when subjects of various kinds are in play, as they



necessarily are in the study of the living, including humans. One kind of subject in particular, the single autonomous individual or person, seems to carry a very great explanatory load when we are called to account for our behaviors and activities. This stark individualism has been pointed out by many critics of modernity, especially of a Western, post-Enlightenment and largely Christian modernity. With the introduction of these highly politicized and polarizing adjectives, it is clear that any treatment of subjects and their objects will be potentially contentious, and that is as it should be. Throughout the book, we will encounter arguments in which the scientific debate is inextricably entangled with political concerns and the foundations of cultural identities. In the final two chapters, some recent perspectives from the enactive tradition in philosophy and science are introduced that may be of service in dealing with this kind of complexity. It is my hope that such debates will be enriched by being approached from a scientific point of view, and through the use of worked scientific examples.

This book makes the case that joint speech merits our attention, that we can learn much by considering how it should be accommodated within existing scientific practices, and that those practices might be expanded or augmented through what we learn. For the window opened by such study does not reveal only a strange form of speech. It provides an empirical access to practices by which order in our lives is created and sustained. It has the potential to lay bare the manner in which several sources of order, normally considered distinct, may overlap and become entangled. These include the regularity of the natural world (natural law), the authority of civil institutions (civil law) and the admonishments of tradition and religion. The study of joint speech is thus not only of interest to one or other academic discipline. It bears consequences for how we conceive of truth, what kind of truths may be arrived at within the scientific domain, and how the authority that comes with knowledge is negotiated politically.

## **Acknowledgements**

My thanks to several friends who contributed through discussion, and engagement with earlier drafts of this text. I have benefitted greatly from many conversations with Marek McGann, Tony Chemero, Stephen Cowley, Nick Campbell, and Jens Edlund. Alexandra Griesser, Paulo De Jesus, Roger Moore, Judit Varga, and Mark James also provided valuable feedback, and opened my eyes to many facets I had missed. The professionalism and vision of colleagues such as Sophie Scott, Kyle

Jasmin, Stephen Connor and Daniel Richardson has greatly advanced my own thinking on joint speech. In the background stand the broad influences of Scott Kelso, David Vernon, and in particular, Victor Langheld, who runs the finest Indian Sculpture Garden in Ireland. Thanks also to the many colleagues and students who have supported discussion on these topics through invited talks, panels, round tables, and the like. Bei Wang, Beatriz Raposo de Medeiros and Carlos Cornejo have played a special role here, as have ongoing discussions with Abeba Birhane, Austin Dwyer and John Francis Leader. Felice Varini kindly allowed reproduction of his works in Figure 17. Let the conversations continue!

## PART ONE

### INTRODUCING JOINT SPEECH AND THE SUBJECT

In this opening section, we introduce the topic of joint speech. In addition to definitions and labels, we illustrate joint speech through several concrete examples. Some readily apparent features of joint speech that give pause for thought are enumerated. These include the absence of an obvious distinction between speakers and listeners, the use of a great deal of repetition, the participatory nature of joint speaking, and the absence of any clear divide between speech and music in such activity. The big question is raised of why there is no science of joint speech.

In order to better understand joint speech on its own terms, it is necessary to recognize the importance of uttering. Joint speech is placed within a continuum extending from interior monologue at one extreme, through the to-and-fro of conversational speech, on to dialogical interaction with call and response in ritual and rite, and culminating in the earnest recitation of solemn texts such as a Credo or an oath of allegiance. The centrality of joint speech in highly valued cultural practices provides an incentive to stand well back from our topic, and to consider its contribution to the broad notion of Logos, understood as a generalized sense of order that finds expression in natural law, in civil law, and in the dictates of religion, tradition and habit.

Modern science has progressed from the dispassionate observation of the stars and planets to the more local and familiar territory of the biosphere in which organisms, singular and plural, from single cells to herds of wildebeests, co-exist in a dynamic negotiation of values and concerns grounded in diverse forms of embodiment and lifeworlds. Among the living, science is forced to confront a multiplicity of perspectives, a task that must still be regarded as work in progress. In the domains of medicine, psychology, and the social sciences, we encounter different kinds of subjects. Joint speech serves to draw our attention to highly

valued human practices that seem to ground collective identity and being for many communities. It challenges us to resist a view of agency and autonomy located only in the individual person, and to see ourselves as essentially and variously collective.

## CHAPTER ONE

### SOME INITIAL OBSERVATIONS

#### **Example 1: Reciting the rosary**

IN THE BEGINNING . . . We are inside a small convent church in County Cork, Ireland. A group of nuns from the Poor Clare Colettine order are leading a recitation of the rosary, an extended form of ritualized prayer, once common in this country, but now slowly dying out. About 20 local lay people are also present. One of the nuns acts as a lead voice. She and her fellow nuns are located on one side of a dividing rail. Most of the nuns are kneeling at small individual benches, but the lead speaker stands in front of a microphone. Her voice is very soft. The lay group is on closely packed chairs facing forward. They are mostly middle-aged or elderly, with some few exceptions. Women outnumber men about 2 to 1. Most people, both nuns and lay, finger rosary beads, to help them keep track of the prayers. The beads, like the prayers, are organized into groups of 10, or decades, each bead corresponding to one recitation of the Hail Mary. Between groups of 10, some isolated prayers are uttered, and the whole suite of decades is bookended by additional prayers spoken together.

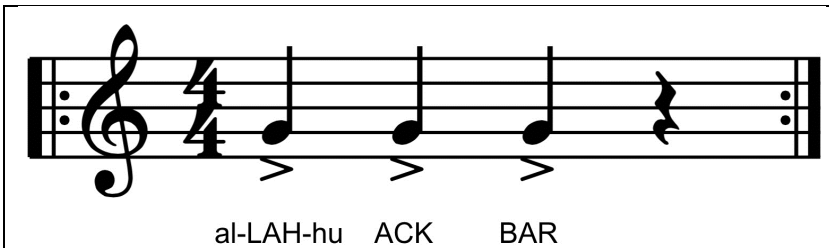
The prayer that is so often repeated has two halves. The first half is recited alone by the leader. The second half is a response, uttered by all present. The other prayers are likewise divided into calls and responses. The prosodic, or musical aspects to the voice, are quite pronounced. Each time through, the words are pronounced with the same slightly lilting melody, not quite sing-song, but not like conversational speech either. Everybody present is very familiar with the practice, and when everybody speaks together, there is a gentle acoustic blur, made all the more indistinct by the reverberant character of the room. Individual words or phrases are hard to hear. Synchronization among participants is loose, allowing some voices to be tracked as individuals. The role of lead speaker is rotated at the beginning of each decade, and the decade is introduced with its title: The First Glorious Mystery: The Resurrection; The Second Glorious Mystery: The Ascension, and so forth. Decades are grouped into sets of 5—the

joyful, sorrowful and glorious mysteries—so that the entire recitation has a complex hierarchical formal structure. Adding to the formal intricacy, prayers in successive decades show an alternation such that what was “call” in one decade is now “response” in another. Each part of each prayer is thus recited by everybody. This leads to a little uncertainty at the start of some mysteries, as not everybody seems to be entirely sure where they are in the structure. But the scaffolding of the collective is enough to establish unity and confidence very quickly.

### **Example 2: Strife at the Al Aqsa mosque**

A MORE VOLATILE SETTING. Leaving the relative calm of the nuns in Cork, we travel now to the plaza outside the Al-Aqsa mosque in Jerusalem. It is February, 2012. There has been friction, and there are wisps of tear gas in the air. Onlookers of many kinds are present, journalists with cameras, tourists, Arabs, Jews. Israeli riot police are also present, all dressed in black uniforms. They group together and an interface forms between civilians and police. There are scuffles. A man is grabbed by the police, and pulled back by his associates. Once he has been recovered, the police and the civilians hesitate. Suddenly a cry goes up from the civilian quarter: “Allahu akbar.” This is the *takbir*, ubiquitous in the Arabic world, misunderstood in the West. It quickly becomes an insistent chant, with three beats stressed out of four: “a-LLAH-hu AK-BAR.”

Now, suddenly, there are two groups, two collective entities, present, where previously there was only one. The riot police are already clearly marked as belonging together. They wear similar uniforms, shields, truncheons; they stand together and move en masse. The civilians were hitherto a colorful mixture, moving like pollen grains on the surface of water in Brownian motion, uncoordinated and various. But when the chant starts, they coalesce, and now there is a second group to stand in opposition to the police. For the brief period that the chant persists, we see protesters versus police. The chant unites, and a collective agent is temporarily brought into being.



**Figure 1: The Takbir**

*“Allahu Akbar” as chanted during one violent conflict outside the Al Aqsa Mosque in Jerusalem.*

The phrase “Allahu Akbar” is famous, or better, infamous, in the West. Its meaning is often rendered as “God is great,” or “God is greater.” It has, unfortunately, become associated with the commission of violent acts, so that to many non-Muslim people, it has become a symbol of violence. Within the Islamic world, it is seen and used entirely differently. It is ubiquitous, so much so that it has its own name, the *takbir*. If one calls out the name of the phrase, those around will respond with synchronized calls of “Allahu Akbar.” The phrase is uttered both individually and collectively, *sotto voce* and out loud, and under many different kinds of circumstance, both delightful and horrific. It does not at all herald or signal violence, but rather functions as an injunction to the pious Muslim to recognize that no matter what he or she is experiencing or doing, it can be relativized by recognizing that God (Allah) is *greater*.

## On observing

Prayer and protest make odd bedfellows. The gravitas of many forms of religious worship is far removed from the chaos of violent protest. Yet these two domains of human activity might be argued to share much in common. At a superficial level, we find the quasi-musical unison chanting of texts whose meaning is completely familiar to the participants. There are associated synchronized gestures (making the sign of the cross, fist-pumping), and there is an awful lot of repetition. These overt similarities might be dismissed as no more than the accidental use of specific forms of cultural practice, creating the illusion of commonality, and hiding fundamental differences in the purposes to which these forms are used. In one case we have a highly formalized ritual, designed and perfected by others, and repeated in more or less invariant form in many places and throughout centuries. In the other we have a highly contingent, improvised expression of frustration and anger directed against a very specific and

tangible target. To link them by virtue of the relatively trivial characteristic of chant might carry no more weight than observing that clothes are worn at both events, and both happen in the afternoon.

But there are advantages to attending to superficial things. That which is on the surface can be observed without further ado. It can be observed by you and by me. That rather obvious characteristic has some benign consequences. It facilitates the path to consensus. If we both observe something, such that we are happy to use the same words to describe it, then we have a starting point for a discussion about the significance of what we have observed. This doesn't stop the merry work of disagreement thereafter, but it does provide a useful starting point. How many arguments go wrong because of the failure to agree on what it is that is under discussion?

The drive to achieve consensus, even limited and partial consensus, underlies science, politics, diplomacy, much of religion, and, in less formal mode, a lot of everyday conversation. In each case, the chances of achieving some kind of consensus are greatly increased if the discussants can demarcate a field of discourse, within which some basic foundations are agreed upon.<sup>1</sup> In the discussion to follow, the subject matter will range over many disciplines, drawing on the everyday experience of familiar practices, but also linking these to scientific, philosophical, and historical arguments. In order to discipline the discussion, it might be wise to stay close to the surface of things, to lean heavily on observations in which we have some confidence, and to return to simple brass tacks whenever possible. In this manner, I hope both the reader and I may emerge unscathed.

The goals of the scientific enterprise will be of special importance in what follows. Science as I understand it starts with observation. In what follows, I will not treat science as if it were a unified enterprise, with each specialization capable of rational alignment and unification with each other. The kinds of observation and argument found in physics bear little relation to those found in biology, psychology, social science, economics, geography, or countless other fields, some of which purists may wish to exclude from the scientific family altogether. The small descriptions

---

<sup>1</sup> Of course consensus is by no means the only, or even main, point of conversation. Fluidity, indeterminateness and ambiguity are essential characteristics of human intercourse, and they make it fun. Sometimes, though, it is good to arrive, even temporarily, at a common point.



provided above of praying nuns and chanting protesters are observations of a sort. They are uncontrolled, to be sure, mere anecdotes, but we will treat them here as observations worth taking somewhat seriously despite this limitation. In coming to understand human practices, careful ethnographic observation is an essential point of departure.

In a rigorous formal framework, individual observations have a determinate form. If we are plotting star positions in the night sky using an agreed coordinate system, two numbers (and perhaps a time stamp) will suffice for each. But, for better or for worse, we will be very far from a rigorous formal framework here. In the journey before us, most observations will demand some contextual embedding. The anthropologist Clifford Geertz introduced the notion of “thick description,” whereby any recorded observation is supported by as much contextual detail and elaboration as possible, thereby facilitating a rich form of interpretation, and holding back from a single, determinate, reading of the data (Geertz, 1973). Within anthropology, such elaborate descriptions become necessary when the objects of study—other people and their practices—are alien to the investigator. Geertz famously used this approach in describing cock fighting in Bali. In what follows, our situation will be somewhat similar, in that we will attempt, at times, to alienate ourselves from our everyday world, to view ourselves from the outside, making the familiar strange, in full knowledge that such an exercise is impossible. The very impossibility of such distance will be a topic we shall have to consider in its own right.

Under these circumstances, the ease with which joint speech can be observed will be a virtue, allowing us to calibrate our observations and to keep two feet firmly on the ground. I will often make use of extended descriptions as starting points and as anchors, tethering the more conceptual arguments to specific instances. In most cases, I will have in mind specific recordings of specific events, so that the details I note are not imaginary ornamentations, but documented features of at least one instance. An archive containing video recordings of specific instances of joint speech is being assembled, and pointers to the particular examples described in this book will be available there.<sup>2</sup>

---

<sup>2</sup> The archive is available at [jointspeech.ucd.ie](http://jointspeech.ucd.ie).

## Joint speech defined

The kind of speech to serve as an empirical anchor here is simple to define. **Joint speech is speech produced by two or more people who utter the same thing at the same time.** I intend “joint speech” as an umbrella term to cover many different kinds of speech produced in very different contexts, yet meeting this minimalist definition. The term “utter” is used, rather than “say,” as joint speech will extend to include many utterances that lie between speech and song. We can identify a small number of subsidiary varieties:

**Choral speech.** This is a genre of performance in which a group, such as a class of children, recites a set text for an audience. The audience will not infrequently be largely composed of relatives of the speakers. Choral speaking competitions are found in many countries. They seem to be particularly popular in Malaysia and South East Asia generally, but are also found in Ireland as a specialty in performance competitions (*Feiseanna*) involving solo recitation, music making, dance, and the like.

**Chant.** The English word *chant* is ambiguous with respect to whether the vocal activity is considered to be speaking or singing. It can be used with equal applicability for the austere plainsong of a group of Benedictine monks or for the raucous hoots of a bunch of soccer fans. This ambiguity will serve us well in what is to come, as the domain of joint speech does not seem to support a categorical distinction between speech and music.

**Synchronous speech.** This is a term I coined to specifically refer to joint speech produced under laboratory circumstances, in which speakers are speaking at the behest of a researcher, and not with their own purposes in mind. The texts employed are usually unfamiliar and of no special significance to the speakers.

Joint speech is found in a wide variety of circumstances, and the few specific varieties noted above do little to circumscribe the activity more generally. But we can use the simplicity of the definition of joint speech in a singular fashion—to pick out discrete and diverse domains of behavior. We can use the definition of joint speech as a kind of lens with which to frame our observations. To see why this might be helpful, consider the task facing the poor anthropologist or behavioral scientist who wishes to study “ritual.” Central examples of ritual are not hard to find—the Roman Catholic mass, or the coronation of a new monarch might provide obvious

and plausible examples. But the borders of ritual are not easy to identify. Does your habit of folding your clothes and placing them on a chair before going to bed count as a ritual? What about tooth brushing? Is a football game a ritual? These are not substantive questions precisely because we lack an agreed definition of ritual. But if we use the definition of joint speech as a means of framing our observations, we find that its simplicity allows us to home in on several singularly important domains of human activity, without tripping ourselves up with such definitional niceties. If we ask “Where do people say the same thing at the same time?” we immediately pick out several familiar domains.

The largest by far is the domain of *prayer*, which for our purposes will generously extend to encompass both *liturgy* and *ritual*, while excluding silent prayer and the supplication of the individual, for these do not fit the empirical definition we are using. When we approach collective prayer, ritual and liturgy with joint speech in our sights, they seem to overlap so much that there is no profit in seeking to carve them into distinct provinces. With that, we are confronted with a widespread human activity that lies at the center of very diverse forms of order. In picking out one or other manner of speaking as prayer, or more specifically as collective prayer, we immediately reveal our own commitments and our own heritage, for that which might plausibly appear to me as prayer will be activity that bears some similarity to the practices I am familiar with, have grown up with, and that are on display around me. Coming, as I do, from an Irish background, the prototype of prayer might well be something like the recitation of the rosary described above. As we move further afield, it will be less clear what counts as prayer and what does not. We might encounter trance-like states induced by twirling, as in the *dhikr* of the Dervishes. This looks little like Roman Catholic prayer. Is it still the same phenomenon? What about the manual rotation of so-called “prayer wheels” by visitors to Tibetan shrines? Or to take an extreme example that clearly illustrates the difficulty of escaping our own grounding, what are we to make of recent reports of chimpanzees in the Republic of Guinea in West Africa, who have been observed hooting and banging rocks against a specific tree, and piling up stones against it. The purpose of the activity is not available to us as human observers. We cannot legitimately claim to understand chimpanzee activity that does not wear its intentions on its sleeve. Yet we find reports in the popular press announcing “Mysterious chimpanzee behaviour could be ‘sacred rituals’ and show that chimps believe in god” (Griffin, 2016). At this point, reason has left the building.

When we start with joint speech as our framing device, we will not delineate the domain of prayer to anybody's satisfaction. But we will be able to group observations together that belong together. We will be able to recognize commonality across many kinds of tradition, despite the fact that the traditions in question employ very different suites of concepts in describing themselves, their activities, and the world. In this sense joint speech studies can inform us about prayer, in a manner analogous to the use of blood pressure measurement to a cardiologist. The circulatory system is complex, and blood pressure measurement provides a very incomplete window into that domain, but it is a useful one, clearly relevant to the functional organization of the system, and it thus provides an empirical anchor to the heart doctor. This may be all the more important when we come to study behaviors and forms of organization that we do not understand, or that are not interpretable with our local suite of concepts, labels, and biases. By focusing on joint speech, we will be looking at activities that might be interpreted as religious ritual, rite, liturgy, or as prayer. The empirical focus provided by the definition of joint speech obviates the need to categorize the activities in too rigid a fashion.

The domain of *protest* also jumps out at us. We find the use of joint speech whenever people gather together to object, demand, or to revolt. There is variation from one situation to another, and such variation will be of interest to us; but it is by virtue of the unison chanting that the domain of protest is approached empirically, helping us to avoid thorny questions about what, exactly, counts as a protest.

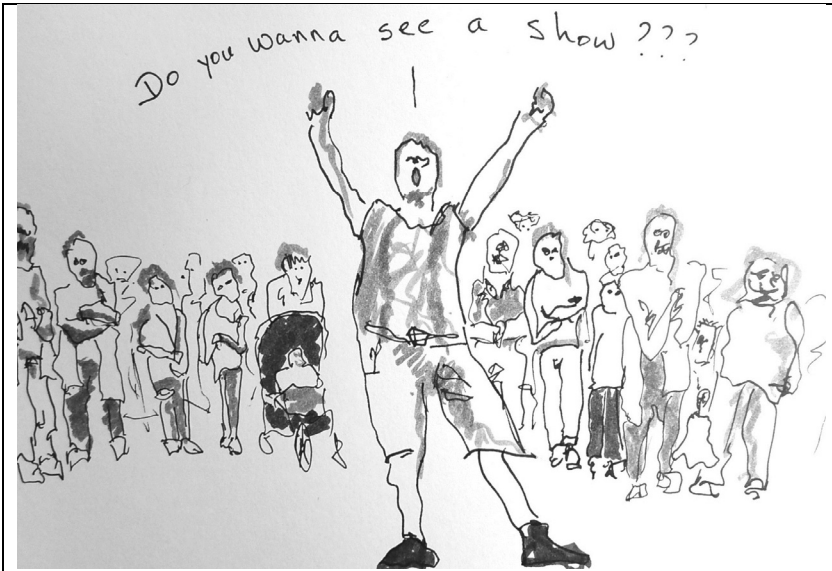
A third domain that we must immediately recognize has, on the face of it, very little in common with either the gravitas of prayer or the urgency of protest. This is the use of chant among supporters of sports teams. Not every sport has a chanting tradition. It is rare in tennis, unheard of in snooker, but completely at home in soccer, ice hockey, baseball, American football, and several other sports. (Interestingly, rugby, which has very much in common with soccer, does not have a chanting tradition, though it does have its own remarkable singing tradition instead.) Despite the profound differences in the type of activity here, it will prove possible to identify characteristics of sports chants that merit consideration along with other forms of joint speech practice.

Prayer, protest and sports chanting are the three biggest domains that joint speech picks out, but there are very many other areas in which joint speech

arises, each time with its own domain-specific characteristics, but also with features that reveal commonalities where they might not be expected. In schools, teachers make use of joint speech in various ways, and as educational practices are not unified, so too there are diverse ways in which joint speech is used. Rote learning is common in classes with young children, and so recitation of multiplication tables can be found in every country. Getting children to speak together is also a simple way to marshal their attention, and skilled teachers of young children will use it as a means to gather and unite the children.

In many countries, religious education is an important part of basic learning, and chanting traditions are used here too, as a way to instill sacred texts indelibly. Madrassas throughout Asia and Africa use chanting as a means of learning the Koran. Hindu sacred texts have long been passed down and protected by chanting, and sutra chanting is part of the everyday experience of the young Buddhist monk. There is thus continuity in many cultures and traditions between the use of joint speech in education and later in rite and ritual.

We will encounter many forms of joint speech in what is to come. Often, these will be vignettes taken from everyday life, unremarkable under most circumstances. Everybody has experience of speaking in unison, even if many of us consciously choose not to join in this or that form. If the reader is alienated by people reciting the Nicene Creed, she may nevertheless assent to joining in with a chorus of Happy Birthday, a small ritual that also harnesses the collective, unified, voice. We may opt not to join in a pledge of allegiance to a secular authority, but when we join the circle of onlookers drawn to a street performer, and we hear an energetic appeal “Do you want to see a show?” we too will probably call out “yes” with one voice, and with that, we are no longer innocent passers-by, but are now part of a committed group of spectators with common focus.



**Figure 2: Joint speech as a technology**

Joint speech obviously serves many ends. As such, it also constitutes a technology that can be adapted to serve many kinds of goals. When the street performer gets the assembled crowd to shout back “yes,” then he knows he has his audience, and they know that they are part of his show. This transient means of gathering attention is common in classrooms, where it assembles the unruly individuals and gives them a common focus. The informal shout of general assent that gathers the crowd on a shopping street has a counterpart in the collective prayer or oath often used to begin a formal meeting, a liturgy, or a ceremony, or in the domestic ritual of reciting Grace Before Meals. Once we speak together, we have common purpose.

### Some perplexing issues

So joint speech is ubiquitous, familiar, almost pedestrian. Yet it harbors many perplexities to be explored in what follows. When examined, these call out for interpretation, but not, I hasten to add, for explanation. An attempt to *explain* any of the following features would be to accommodate them within an agreed interpretive framework, to assimilate them to the known and secure. I believe that joint speech will resist such a comfortable exercise, and will demand rather more of us. We start by scratching the surface of the phenomenon, expecting perhaps to uncover a novel genre,

style, or cultural practice, one that might provide a pleasant distraction. But what we find is something vastly richer, and more challenging.

Here then are just a few features of joint speech that might give us pause for thought.

- In joint speech, there is no distinction between speaker and listener. This is rather obvious. All participants are engaged in something that is both, or neither of these. To say that prayers are addressed to a transcendent deity does not change anything. A Catholic God will, by now, be completely familiar with the text of the rosary. The protesters may be addressing a political establishment, but most of the time, those addressed are not present. There is no news value in what is said, but it must be said anyway.
- Repetition, repetition, and more repetition. Whether on the street, the terraces, or the church pew, repetition is a canonical feature of joint speech; so much so, that when it is absent, as in the collective swearing of an oath of allegiance, it is the absence itself that is noteworthy. The rosary beads of our opening scene are repeated in Christian Orthodox, Hindu, Moslem, Sikh, Jain, Bahá'í, and Buddhist practices of prayer.
- Performativity. The text of the rosary may be known, but that is irrelevant. It must be *uttered*. The necessity of actually uttering some phrases is well known from speech act theory (Austin, 1962). But within that framework, performatives such as “I dub thee a knight” or “I do” (at a wedding) are relatively rare. They accomplish something only under very specific circumstances, and that something is typically singular. Having married a person, it is not really possible to marry them again (at least not right away). Joint speech is performative, but in a rather different sense. We might speak of *enacting* rather than accomplishing. The vocabulary of enaction will be of use to us as we proceed.
- Musicality. Language (including speech) and music (including both chant and song) bear interesting relations to each other, and much effort has been spent in considering both their commonalities and differences. However when we begin to observe joint speech, any firm boundary between the two becomes invisible. In repetition, strong syllables are exaggerated; intervals are regularized; strong-weak alternations are enhanced; gestures such as fist pumping or clapping are frequently used. All this leads to an enhanced rhythm. Repetition tends to turn intonation patterns into melodies too. The

prosody of joint speech, and the inextricable mingling of the fields of speech and music will be of great interest in what follows. Our definition of joint speech will have to extend to include unison singing, as well as speaking.

A final point to note about joint speech is that the practices we have identified, and others we could pick out by using joint speech to frame our observations, are all accorded a great deal of subjective significance by those who take part in them. The importance of prayer needs no argument. The urgency of protest is evident. The enthusiastic enactment of collective identity on the football terraces makes patent its charms for practitioners. Along with these canonical examples, we might note the solemnity accorded public group recitation of oaths of allegiance and fealty integrated into secular ceremonies. Clearly, this slightly odd business of saying the same thing at the same time is of some considerable importance, and there is ample motivation to seek to understand such practices in a manner that goes beyond the concerns of any single domain.

But now we come to the most perplexing feature of joint speech: it has not been made an object of concerted empirical inquiry at all. There is very little scientific work done on any aspect of joint speaking. As a topic in its own right, it seems to be invisible to those who study speech and language, and to students of human behavior. There are, of course, specialist and scholarly works that approach musical questions such as the history of plainsong and Gregorian chant, or that address liturgical niceties such as the respective roles of priests and congregations in mass. We will even find encyclopedic coverage of the rich and raucous world of football chants. There has been passing acknowledgement within ritual studies of the importance of collective speech and associated gestures in the specific context of religious ritual. There has been a small amount of documentation of protest chanting in specific situations, such as during the tragically misnamed Arab Spring of 2011 (Moghith, 2014). What is missing is the thematization of joint speech itself.

Joint speech is absent from linguistics. Speech is not the same thing as language, and we will have cause to consider features of speech that have no counterpart in language, conventionally defined. The scientific study of language has a history of focusing on the encoding and transmission of messages, abstracting rapidly away from the messy business of shouting, whispering, cajoling, and imploring in specific communicative contexts, to the more rarified and untethered domains of syntax, semantics and



phonology. This has made some aspects of the spoken voice less visible that they might have been. In recent decades, some researchers have begun to probe vocal behaviors more attentively, paying attention to those elements that do not make it onto the page in writing, including such neglected particles as uhms, ah-has, grunts, and silences. Where once linguists approached the sounds of speech to try to recognize the ghosts of letters and words, more recently phoneticians have begun to study the melody and rhythm of speech, the placement and duration of pauses, and the rich variation found in voice quality (Wennerstrom, 2001). The intimate relations between speech and gestures are starting to be traced (Wagner et al., 2014). Joint speech adds a great deal of additional material for study, though I would suggest that it must be studied on its own terms, and not merely as the voice of the individual, replicated many times over.

Joint speech is absent from behavioral and movement sciences. The vast majority of behavioral science looks at the actions of distinct singular persons. The form of any kind of skilled movement bears the signature of the individual. Though we may all reach a similar level of proficiency in writing, in walking, and in speaking, the manner in which we do so marks us out as unique and distinct, and the bodily patterns we exhibit as we perform similar tasks all speak of our individual identity, our accent, our uniqueness. Babies almost all learn to walk, but some do so by crawling, some by bum shuffling, and some by observing quietly. More recently, however, researchers have begun to pay keen attention to important aspects of behavior that are not captured by studying individuals. The shoaling of fish, flocking of birds, even the collective behavior of crowds in various physical environments are all now attracting attention (Vicsek and Zafeiris, 2012). But joint speech has so far been missing. Among the many questions we can ask, we might consider why joint speech has seemed to offer so little to observers of human behavior.

Joint speech is absent everywhere. A quick search on Google Scholar, the search engine of choice for scientific and academic publications broadly considered, reveals little. Part of the problem is terminological. In the absence of an established field of study, I have introduced the term Joint Speech, and I confess it was I also who introduced the term Synchronous Speech. Choral Speech is of greater antiquity. One might also look for Unison Speech. Combine all of these search terms, including both “speech” and “speaking” as variants, and I can find no more than about two or three thousand works, many of them accidental catches. The greatest number of published works in the field address “choral speaking,”

and most of those belong to the slightly quaint field of elocution, whereby school children are taught to recite entertaining verses with polished pronunciation to the delight of appreciative parents.

By way of comparison, we might look for scholarly works on the slightly odd phenomenon of glossolalia, or speaking in tongues. This is a vocal behavior found within some evangelical religious traditions, and it is especially prevalent in Pentecostal congregations. It is, by all accounts, a very moving experience to utter syllables without a determinate meaning. Believers generally attribute the source of their utterances to the Holy Spirit. Google Scholar produces over eleven thousand scholarly publications that address this specific behavior. They include many works in the domain of cultural studies, comparative anthropology, ritual studies, and theology. There are neuroscientific studies, psychological studies, consideration of possible relations to psychopathologies, and even phonetic studies. There is work on glossolalia as a learned behavior and as a form of possession. Relations between glossolalia and personality types are explored. Glossolalia is a fine topic for research apparently. Yet when we do some mundane counting, it is clear that instances of speaking in tongues are clearly outnumbered by instances of joint speech, not by a hundred to one, or a thousand to one, but by literally billions to one, for joint speech seems to occur in all societies, in many different domains, and it is difficult to conceive of a vocal individual who has not spoken in unison with another at any point, while few of us, with the respectful exception of Pentecostal congregation members, will have spoken in tongues.

This then is the conundrum I wish to look into in depth. Why has joint speech remained invisible, despite the rather obvious facts that it has quite distinct characteristics, is accorded the greatest significance by practitioners, and is easy to observe. Is this neglect? Is there perhaps nothing to see in a group of people speaking in unison that cannot be found in the voices of individuals? In many respects, the apparent invisibility of joint speech may be its most interesting feature. As we shall see, it is neither difficult, nor unrewarding, to study joint speech as a scientist. There are many aspects worthy of consideration, and with minimal effort, we can generate scientific findings that speak to linguists, behavior and movement scientists, neuroscientists, social psychologists, and many more. But the absence of such work speaks of something even more important.

## Of subjects

The scientific tradition we value has its own history. It did not spring into the world fully fledged. If we squint a little, we might detect the modern scientific viewpoint coming into being first in the domain of astronomy, as the challenge of interpreting the motions of the planets, the moon and the sun from an earth-bound perspective was addressed, and ultimately solved. Objective science at its best has led us to learn how to think of our position on Earth, in a vast universe, most of which is alien, inanimate, and remote. As we move nearer to our terrestrial home, the kinds of studies we now understand as belonging to physics and chemistry were the next to emerge, and with Newton's magnificent construction of a theory of mechanical motion, it was possible to generalize from the movements of bodies close to hand all the way to the impersonal and imperious passage of the planets in the night sky. But the application of the scientific method to the goings on of living beings took a while. A scientific biology did not really appear until the beginning of the 19th Century. Scientific psychology emerged later in the same century; the application of scientific thinking to societies and groups of humans did not begin until the 20th Century. The social sciences might reasonably be considered to be still grappling with the task of finding basic concepts and methods that can garner widespread consensus.

As science has turned towards the living, and ultimately towards our own selves, the cool disinterested gaze of the observer has been challenged. With the study of the living, it becomes necessary to recognize and consider the perspectives of the living themselves. Living beings are subjects, not mere objects. They have *perspectives*. Things *matter* to them. The notion that science might provide a single God's eye view from nowhere, with no reference to value or to the perspective of a subject, now appears somewhat naïve (Rorty, 1979; Nagel, 1989). The inestimable profit accrued from the application of the scientific method to inanimate matter makes it inevitable, even obligatory, that we should apply those same methods to the goings on of the living, and to human affairs. But where astronomy can get by just fine in an objective key, any science of the living must grapple with subjectivities of many kinds. As soon as we must appeal to any notion of *function*, then there is a subject lurking behind that appeal. The healthy living body is a subject for whom a beating heart can perform a function. Value-laden battles are fought literally under our skin as we speak of pathogens and anti-bodies, a view predicated upon the body as a domain for whom encounters with microbes

are meaningful. We cry out for science to deliver results we can use in medicine, in education, in regulating our own conduct as individuals and as groups, but this kind of science cannot be done as if the entities involved were mere objects and their interactions were free of value and significance, at least to the entities themselves.

The study of joint speech opens up new opportunities here. The activities we have surveyed above are steeped in values. The values are collective, and the subjects associated with them are collective subjects. Here we begin to see why there is such a remarkable absence of empirical work taking joint speech as its topic. We have not yet developed a language with which such collective values and collective subjects can be adequately addressed. Science in a simplistic objective key recognizes no subjects. The science of the living, or biology, in common with the psychological and the social sciences, has yet to find a way to rise to the challenge of integrating the competing and conflicting perspectives of multiple subjects, each with their own set of values. The old fashioned idea that science does not traffic in values has had its day. Science, when turned to goings on within the biosphere, within society, and by and for humans, has no option but to carefully negotiate the presence of many actors, many kinds of value, and multiple perspectives. This may be old hat to social scientists, but there is work to be done in establishing continuity between, and conversation among, the social sciences, the human sciences, the sciences of life, and the so-called hard sciences. As our gaze is drawn to those practices in which joint speech regularly occurs, we must confront the inextricable mingling of political and cultural concerns along with our scientific practices of observation and measurement.

So when we undertake the study of joint speech, we run the risk, and encounter the opportunity, of pushing empirical science where it currently does not go. We will need to acknowledge various kinds of subjects that arise through the collective activity of many kinds of groups. We will have to do so, in self-conscious awareness of our own limitations, of our own biases, our own grounding. The challenge joint speech presents is not that of an indecipherable object of study. As an object of study, joint speech is fascinating, rich, and ripe for the picking. It is its complement, the *subject* of such speaking, that will throw up the greatest challenges in what follows.

There is no way to address these topics without venturing waist high into contested territory. In so doing, my own shortcomings as an observer and

interpreter will become apparent. Joining the dots across radically different disciplines is a challenge, and within the human sciences, it should be recognized as an unavoidable challenge. The strategy to be adopted here is to constantly return to the surface, to observe together specific examples, and to use these to anchor the discussion. Joint speech is an empirical locus that can do service to philosophers as well as to scientists, and can inform the interested non-specialist too, providing footholds where discussion becomes difficult and fractious, and providing rich material to try out new ideas about how science should be done.

