Many confident statements have been made about the central role that coordination plays in social human interaction. For example:

“To successfully interact with others, people automatically mimic their actions and feelings.” ¹
“Coordinated behavior patterns are one of the pillars of social interaction”. ²
“...the mechanisms of coordination that are fundamental to successful interactions” ³
“Successful interpersonal communication depends to a large extent upon the exchange of nonverbal information” ⁴
“The temporal coordination of interpersonal behavior is a foundation for effective joint action” ⁵

Q. In what ways might we investigate the role of synchrony and coordination in human communication and interaction? In choosing what to study, do we face the situation that the more designs and phenomena allow precise measurement, the less the study will about social interaction, and the more the study is about social interaction, the less precise it will be?

Background:

The experimental work connected with the above declarations, often involves rather artificial tasks. For example, people making simply oscillatory movements: tapping to maintain a beat⁶; repetitive arm curls while holding weighs⁷; or people making judgments about others who are carrying out such movements, e.g., waving in or out of synchrony⁸. Although the data from these procedures are relatively straightforward to analyse (although see⁹), they are somewhat removed from realistic scenarios, this begs the question of whether the evidence obtained is germane to human interaction in the broad.

Other types of study, have used more naturalistic settings; attempting to assess coordination between actors, and then link this to factors likely to influence social interaction. For instance, a role-played job interview with mixed/sex and same/sex interactions⁴; free interacting sessions in noisy/quiet environments⁷; friend and non-friend dyads playing a computer game in neutral/conflict situations⁸. Such studies need to first extract some dynamic property relevant to social interaction; submit these data to an analysis of synchrony/coordination, and then determine whether the observed level of synchrony differs from that of sequences that are not influence by interaction. Properties that have been measured include such things as global motion⁸; head motion⁶; and hand motion⁷. Methods of establishing coordination include cross-correlation; recurrence analysis and cross-spectral coherence; with ways for comparing against ‘non-interaction’ sequences often involving the construction of surrogate datasets. These measures, especially cross-correlation, need to take into account auto-correlation, see⁹. Moreover, there is the issue of causation, and testing what causes what may be tricky given the global nature of some measures, e.g., overall energy. Even when some unambiguous dynamic relationship between people has been established, there is still the issue of what it means and what function it has. This has led to another set of studies that have examined the timing and occurrence of events that appear to be more meaningful, e.g., higher-level behavioural events such as smiles, gestures, gazes, and especially linguistic phenomena.

Accounts that attempt to describe and explain the dynamics of conversation are typically are couched in terms of mental constructs such as representation and intention. For example, in the view of Clark¹⁰, interlocutors coordinate by making inferences about the intentions underlying their
partners’ behaviour. On this account, coordination is successful if interlocutors can develop mutual beliefs about their intentions; the alignment of these beliefs is facilitated by conventions, e.g., norms for the allocation of turns, etc, that map intention to behaviour. Similar accounts have been formulated based on the notion that the processes acting of representations underlie coordination and, ultimately, mutual understanding between interlocutors\textsuperscript{11}. Here, it is argued that alignment between interlocutors occurs because linguistic representations at different levels (e.g., phonology, syntax, semantics) prime each other and these representations are shared across production and comprehension. Evidence for this type of alignment often involves assessing the degree to which interlocutors converge (e.g., speak about things in the same way). Convergence is argued to occur at multiple levels, with the evidence tending to become more descriptive at more abstract levels. Even at the phonological level, the evidence for convergence tends to be variable, as a recent meta-analysis makes clear\textsuperscript{12}.

So, how to progress?

References