1   Introduction: why study conversation?

This chapter shows the importance of studying conversation as a route into understanding language in social life. As an introduction to the chapters that follow, it establishes the 'Two Things': the two fundamental things at the core of conversation analysis (CA): action and sequence. We examine some basic linguistic conceptions of the purpose of language, the often indirect relationships between grammatical forms and functions, and the role that 'meaning' and 'context' have played in the investigation of language, within the domains of semantics and pragmatics. Understanding how actions are accomplished collaboratively across sequences of talk provides insights into the basic infrastucture of interaction that may be overlooked in the face of the structural diversity of, and constant evolution in, the languages of the world. We explore areas of commonality and difference between CA and other domains of work within linguistics. This is done initially through an overview of three dominant theories within pragmatics, Speech Act Theory, Gricean implicature, and Relevance Theory, showing something of how some of the phenomena examined in these approaches are treated in the data of CA. We then examine more overtly observational approaches, such as sociolinguistics, interactional linguistics, anthropology and discourse analysis, to show the distinctive contribution CA brings to work on language in interaction: one that goes far beyond the traditional domains of linguistic study. This Introduction ends with an overview of the chapters to follow.

We live our lives in conversation; between the first 'hello' and the last 'goodbye', conversation is where the world's business gets done. Each of us owes our very existence, at least in part, to conversation; we conduct our lives through it, building families, societies and civilisations. Yet the means by which this is done is anything but obvious. This book is an introduction to the study of conversation through the methods and findings of conversation analysis (CA), the domain that has done more than any other to examine interaction, that is, action between people.

Language – at the meeting point of biology and culture – has been the object of intellectual inquiry for centuries, and long regarded as the core of what it is to be human; the investigation of language structure is a basic project in the cognitive sciences. However, only in the last half-century has systematic attention been given to the domain of interaction – where language may be the central component, but not the exclusive one.

In taking interaction as its focus, this book seeks to investigate the communicative and cultural constraints shaping language as they intersect with the cognitive. It takes the stance that, to establish what it is to be human, what happens *between* minds – the visible work done by participants in interaction – is fundamental to finding out what is *in* them. We start by establishing the twin foundations of CA: **action** and **sequence**, and how they promise to illuminate some of the central concerns in linguistics. Through the lens of these, we examine a number of traditional linguistic domains to offer some of the insights CA has

made into some long-standing linguistic conundrums concerning the meaning of utterances. In doing so, we not only show what linguistics has to gain from an apprehension of both action and sequence, but also introduce the analytic themes to be pursued in the following chapters.

1.1   The basics: the 'Two Things'

The 'Two Things' game invites us to identify the two fundamental things about any domain.[1] For CA, the two things from which all else follows are *action* – broadly, the things we do with words[2] – and *sequence* – 'a course of action implemented through talk' (Schegloff, 2007a:9). For those more used to dealing with sentences, utterances, meaning and grammar, the terms represent a wholesale methodological tilt of a familiar planet, linguistics, on its axis, and it may not be immediately apparent how either figures in language use. However, the importance of both goes back to the very origins of human evolution.

While the dating of the origins of language is a matter of some dispute (see Dediu and Levinson, 2013, for a reassessment of the usually quoted 50,000–100,000 years to half a million years), what is beyond doubt is that there is evidence of cooperation among our earliest human ancestors, *Homo habilis*, around 2 million years ago. In other words, joint action in the form of cooperation and coordination has been central in the development of humankind. In addition, around seven thousand languages, at a rough estimate, have evolved and are in present-day use; and the study of this diversity and its origins provides linguistics with some of its fundamental and motivating questions regarding the basis for linguistic structure and the nature of its biological and cultural underpinnings.

In seeking to understand this diversity, many have recognised the origins of language change in language use (see, e.g., Hopper, 1987; Lehmann, [1982]1995; Croft, 2000; Coussé and Mengden, 2014; Bybee, 2015). As Evans and Levinson put it: 'most linguistic diversity is the product of historical cultural evolution operating on relatively independent traits' (2009:444). Thus, the examination of linguistic structure reveals 'general cognitive abilities: the importance of repetition in the entrenchment of neuromotor patterns, the use of similarity in categorization, and the construction of generalizations across similar patterns' (Bybee, 2006:730; see also Edelman, 1992, and Hurford, 2007).

There are, of course, uses of language that are not embodied in interaction – jotting down a shopping list, reading a novel or working on a computer – but overwhelmingly, we encounter language, and are socialised, in interaction, and specifically in that particular form of interaction that we recognise as ordinary

[1] This proposes that for any subject, 'there are only two things you need to know. Everything else is the application of those two things, or just not important', e.g. trading in stocks and shares: '1. Buy low 2. Sell high'; acting on stage: 1.'Don't forget your lines' 2. 'Don't run into the set.' See Glen Whitman, 'The Two Things' website, currently at: www.csun.edu/~dgw61315/thetwothings.html.
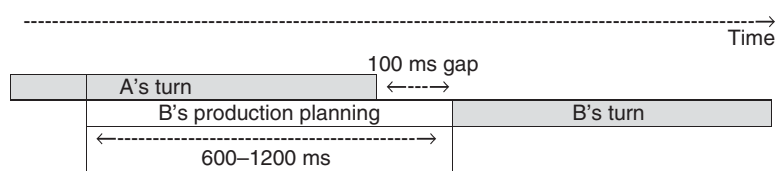[2] To paraphrase Austin (1962).

Figure 1.1 *Overlap of comprehension and production processes in conversation (based on Levinson, 2013a:104)*

conversation.[3] This is equally the case in the phylogenetic development of language down the ages as in the ontogenetic development of the individual. A major contribution to work on the emergence of linguistic structures has been that which examines the discourse basis of various grammaticalisation traits, whether features such as grammatical transitivity (Hopper and Thompson, 1980), lexical categories (Hopper and Thompson, 1984), syntactic change (Givón, 2008, Traugott, 2010) and phonology (Bybee, 2001).

Alongside the interest in the interactional foundations of language evolution and structural diversity, work in psycholinguistics (Clark, 1996) and formal linguistics (e.g. Traum, 1994; Ginzburg, 2012; Ginzburg and Poesio, 2015) has also sought to ground investigation of language in its interactional home base. As such work has recognised, actions – and specifically linguistic actions, such as requesting, inviting, complimenting, complaining, agreeing, disagreeing and so forth – are not unilateral, but jointly and collaboratively achieved. Moreover, as we shall see in the course of this book, this applies as much to actions that, on the face of it, appear to be unilateral, such as referring or informing.

While the value of investigating language as action is thus recognised in many domains of linguistic research, less so is the means by which action is implemented: the sequence. In its focus on how actions are implemented across sequences, CA is committed to studying the spontaneous online production and understanding of language *in time*. One of the most striking facts about the temporal production of turns-at-talk is that while it takes over 600 milliseconds to plan and produce the shortest turn in conversation (Levelt, 1989), on average, and depending on the particular language, gaps between conversational turns are around 200 milliseconds (de Ruiter et al., 2006, Stivers et al., 2009); see Figure 1.1.

There thus has to be an element of linguistic 'double-tasking' in comprehension and production processes. As Levinson notes, conversational participants

> must have parsed what they have heard and understood its grammar well enough to predict both the content and its structure, so that they can predict when it will come to an end (otherwise their response may come too early or too late) . . . action ascription involves numerous dimensions . . . so it would seem to

---

[3] For talk that is not conversational, and a product of specific, standardly work-based, contexts, such as the medical encounter or courtroom exchanges – so-called institutional talk – see Drew and Heritage (1992) and Heritage and Clayman (2010). See also Chapter 4 for how the turn-taking system for ordinary conversation constitutes the baseline for such institutional talk.

> be a much more complex and indeterminate process than decoding the structure
> and content of the turn. That is the miracle . . .   (2013a:103–4)

The implications for interaction across languages of diverse structures, whether left- or right-branching, are profound. Prosodic, syntactic and pragmatic signalling of turn completion or incompletion is directly motivated by the turn-taking system, so that, for example, the English relative clause structure in the turn 'I am reading the book which I gave you' is potentially more vulnerable to overlap than the equivalent clause in the comparable Dravidian or Japanese turn, glossable as 'The I to-you given book I am reading' (Levinson, 1983:365). In the development of the discipline, and in the body of research to date, the CA focus has been just this miracle: how actions are implemented and recognised in talk-in-interaction – or, in its abbreviated form, 'talk'[4] – a term now preferred over 'conversation' as the more general designation for our interactions through language.

The cross-linguistic study of coordinated action is in its early days relative to the long-established research programme in modern linguistics. Moreover, evidently the bottom-up, rigorously empirical working methods of CA are not conducive to making top-down generalisations. However, the search for universals initiated by Chomsky (1957, 1965) and, from another perspective, by Greenberg (1963, 1966), and still subject to vigorous debate (see Evans and Levinson, 2009; and Levinson and Evans, 2010) resulting in the more recent proposal that recursion is 'the only uniquely human component of the faculty of language' (Hauser et al., 2002:1569)[5] focuses in no small part on issues of methodology, interpretation and standards of evidence in linguistics (see, e.g., Everett, 2005, 2009; Jackendoff and Pinker, 2005, and the responses in Nevins et al., 2009). In contrast, the conversation-analytic focus on participants' own displayed understandings has delivered incrementally. However, as we shall see, these bottom-up methods have already yielded enough evidence to suggest that in the face of all the structural variation and diversity across the languages of the world, elements of the procedural infrastructure of interaction (Schegloff, 1992b:1338) studied by CA are indeed universal. So turn-taking, the organisation of sequences, the conversational preference for particular actions and the organisation of repair mechanisms in talk – all part of that procedural infrastructure – are proving to be empirically robust across languages and language groups. As Levinson observes:

> language is held to be essentially universal, whereas language use is thought
> to be more open to cultural influences. But the reverse may in fact be far more
> plausible: there is obvious cultural codification of many aspects of language
> from phoneme to syntactic construction, whereas the uncodified, low-level
> background of usage principles or strategies may be fundamentally culture-
> independent . . . Underlying presumptions, heuristics and principles of usage
> may be more immune to cultural influence simply because they are

---

[4] Schegloff notes that none of the research on embodiment and bodily conduct has undermined any
    of the findings established on the basis of talk alone (2009:360).
[5] We return to this issue in Chapter 3.

> prerequisites for the system to work at all, preconditions even for learning
> language.   (2000:xiv)

The chapters that follow explore the implications of the methodological tilt towards 'action' in 'sequence', starting in this Introduction by examining some of the foundational work in CA and what it has to offer linguistics. It first examines some approaches to language use and the search for meaning within semantics, pragmatics and sociolinguistics, and then briefly examines the domains of interactional linguistics, linguistic anthropology and discourse analysis as the areas of investigation with the greatest perceived overlap with CA. However, despite areas of common interest, there are also areas that are methodologically and perspectivally distinct; this stakes out the basic territory.

1.2   The view from linguistics

1.2.1   The search for meaning

> All sciences search for underlying regularities – that's the game, and there is
> no branch of linguistics . . . that is not a player . . . The art is to find the highest
> level generalization that still has empirical 'bite'.   (Evans and Levinson,
> 2009:475)

The paper in the *Proceedings of the National Academy of Sciences* puts it crisply: 'language has two functions: to convey information and to negotiate the type of relationship holding between speaker and hearer' (Pinker et al., 2008:833).[6] In this, it virtually echoes the perspective of Malinowski a century ago, who in coining the phrase 'phatic communion'[7] thereby gave students of language licence to dismiss certain things done in interaction as essentially social and so unworthy of consideration by students of language:

> Are words in Phatic Communion used primarily to convey meaning, the
> meaning which is symbolically theirs? Certainly not! They fulfil a social
> function and that is their principal aim, but they are neither the result of
> intellectual reflection, nor do they necessarily arouse reflection in the
> listener . . . we may say that language does not function here as a means of
> transmission of thought.   (1923:315)

This emphasis on language as essentially 'transmission of thought', its object 'to convey meaning', is preserved in the traditional division of labour within linguistic study.[8] Here, in very broad terms, the study of word and sentence meaning,

---

[6] In his best-selling popular linguistics book, Pinker puts it even more simply: 'This is the essence of the language instinct: language conveys news' (1994: 83).

[7] 'phatic communion . . . a type of speech in which ties of union are created by a mere exchange of words' (Malinowski, 1923:315).

[8] One prominent pragmatic approach, namely Relevance Theory, does consider so-called phatic communication within the scope of cogntive pragmatic theory (see Žegarac and Clark, 1999) but does not question the essential distinction between the so-called phatic and non-phatic.

largely the preserve of semantics, overlaps with the domain of pragmatics. This takes utterance meaning – that is, the meaning of a sentence in its context (Bar-Hillel, 1970; see also Levinson, 1983: 18–19) as the object of investigation. Within the study of pragmatics, three theoretical perspectives in particular have focused on utterance meaning: Speech Act Theory, Grice's theory of implicature, and Relevance Theory. The last – and indeed most recent – of these three, Relevance Theory, has, of all pragmatic approaches, attempted to address the issue of context and how it figures in utterance interpretation. In its concern with context, it shares a focus, if not a methodology, with sociolinguistics, the empirical orientation of which would suggest it has common cause with CA. Moreover, both interactional linguistics and discourse analysis share some, but not all, of the aims and methods of CA. The following sections briefly examine each of these domains of study in turn to establish the similarities and differences between these approaches and CA; they show why, when it comes to interaction, CA puts action and sequence at the heart of its investigations.

On semantic meaning: stability in action

The linguistic emphasis on language as information transfer, embodied in what Reddy (1979) identifies as a conduit metaphor (cf. Malinowski's 'words ... to convey meaning'), suggests that meaning – encoded and then decoded in the act of communication – is linguistic 'cargo'. However, it is evident that as soon as we examine interaction, a conception of 'meaning' may be enriched by a consideration of both action and sequence. Take, in the first instance, a simple example, grounded firmly in the realm of the so-called phatic: *thank you*.

The fact that meaning does not necessarily map straightforwardly onto use is evident when we consider the meaning of *thank you* in French. Standardly, this is taken to be *merci* – and, indeed, in the context of, for example, accepting a gift, this equivalence holds. However, in one everyday context, it is clear that, in actual fact, *thank you* in English is used in just the opposite way to *merci* in French, and that is in response to an offer. Where a standalone *thank you* accepts an offer, a standalone *merci* rejects one: the actions implemented by these apparent semantic equivalents are thus here entirely contrastive. Here, a search for meaning turns up apparent equivalents, whereas an investigation of action reveals them, in this interactional context, to be sharply divergent.

So understanding that *thank you* might be appropriate to accept a gift, but not, in English, to refuse a drink, depends upon our recognition of what the prior turn was doing. Moreover, with its interactional production, the sequential properties of *thank you* become evident, in that the prosody of *thank you* intrinsically anchors it in a specific sequential position. That is, while stress on *thank* announces this action as initiating thanks, stress on *you* announces it as *reciprocating* thanks. *Thank **you*** implicitly proposes itself as responsive to a prior expression of thanks.

Action, then, is implemented across sequences. Furthermore, if an example such as *thank you* appears mundane and inconsequential – indeed, to be dismissed as merely phatic – consider this: one contributing factor in the world's biggest airline disaster to date, at Tenerife airport in 1977, was a misunderstanding of what the apparently mundane word *Okay* was doing (Roitsch et al., 1977). The message from the cockpit of a KLM plane to the control tower, 'we are now uh-takin' off' or 'at take-off' (the recording is unclear) is met by 'Okay' and then a pause of nearly two seconds. The next portion of the utterance is obscured for the KLM pilot because of radio interference. Here, it subsequently emerged that the control tower was using 'Okay' to acknowledge the prior talk – as a receipt token (on which, see Schegloff, 1982). The pilot, however, taking 'Okay' not simply as a receipt but to authorise take-off duly did so, unable, in thick fog, to see the Pan Am plane in his path. Five hundred and eighty-three people lost their lives in the ensuing collision.[9]

We use the same resources to implement actions across sequences, whether apparently insignificant or hugely momentous – and it is this consequentiality (or, rather, con-sequentiality) of such communicative actions that an appeal to meaning does not wholly capture, even in the case where the meanings of *thank you* or *OK* are intuitively accessible. For many lexical items, intuitions with respect to meaning are reasonably straightforwardly accessed, and indeed fairly malleable; so, as Heritage notes, 'the typification "drink" may be revised towards a more 'fringe' meaning, if when offered "a drink", your host is boiling a kettle' (2011:264).[10] In this instance, context clearly picks out the typification, just as a head nod might either be – according to context – accepting an offer of a drink or buying a Ming vase worth millions (on the latter, see Heath, 2013). But how to understand context when, in the case of linguistic objects (in the most general sense of the term), the semantic core itself may not be easily accessible? The meaning of *drink* or, for that matter, *thank you* or *okay* may be straightforwardly and readily available to intuition, but this is by no means always the case. It is at this point that the analytic relevance, not only of action, but of sequences of action, becomes apparent, for the apprehension of both meaning and context.

There are clearly cases where specifying the meaning of a linguistic object is not straightforward. Take, for example, the commonly used English particle *actually*. A search for 'high level generalisation', as noted by Evans and Levinson, clearly needs to account for something so recurrently used in conversation; and yet, it does not follow that recurrent use can necessarily lead a

---

[9] *Okay* was, as a consequence, dropped from the authorised standard phrases used in air traffic communications. For conversational uses of *okay*, see Beach (1993).

[10] As Johnson-Laird observes, nouns are in fact more like pronouns than is commonly recognised. As with Heritage's example of *drink*, Johnson-Laird illustrates how, at utterance level, context picks out salient aspects of any given object on the example of the lexical item *tomato* where different features are selected by utterance context – in (a) its spherical shape, (b) its colour and (c) its squishiness:
  (a) The tomato rolled across the floor
  (b) The sun was a ripe tomato
  (c) He accidentally sat on a tomato   (1987:197)

native speaker to formulate what *actually* means.[11] However, it is here that starting, not from the generalisation but from the 'empirical bite' – examining a linguistic object on various occasions of its interactional use – may provide some analytic yield. Of course, 'bite' had been effectively ruled out of the game by Chomsky in setting out the main aims of linguistics in his distinction between 'competence' and 'performance':

> A record of natural speech will show numerous false starts, deviations from rules, changes of plan in mid-course, and so on. The problem for the linguist, as well as for the child learning the language, is to determine from the data of performance the underlying system of rules that has been mastered by the speaker-hearer that he puts to use in actual performance.   (1965:4)

However, excluding any investigation of 'the data of performance' because of its apparently 'degenerate quality and narrowly limited extent' (p. 58), and seeking underlying regularities from idealised and abstracted linguistic data risks ruling out of the game just the 'record of natural speech', in Chomsky's words, which may be necessary to the investigation – material such as the following:[12]

(1)   (Clift, 2001:277; H88:U:2:2)
      (L=Lesley, K=Kevin. Gordon is L's son, who has just done a driving test (l. 2);
      Katherine her daughter, who is currently away at university.)

```
1L    [hYe:s. Oh: shame.h.hhhh Gordon didn't pass his
2     test I'm afraid,h=
3K    =Oh dear
4L    .k.tch He's goin- (.) Well.hh u-he was hoping tih get
5     it (0.2) in: uh in the summer but u (.) they're getting
6     very booked up so I don't know if he'll even: get it in
7     the:n.h
8     (1.1)
9K    Yes I: ah: no doubt he's back e (.)t uh
10    (0.5)
11L   .hhhh Yes. We're going up- (.) we:ll- (.) we're get(0.2)
12→   actually it's g'nna be a rather busy Ju:ne, Kathrine's
13    home f'three weeke:n:ds. As it happens people're coming
14    do:wn'n c'n bring'er down which is rather nice,
15    (1.2)
16L   which e-aa::: so we're rather looking forward t'that,hh
17    (1.5)
18L   hA[n:
19K      [Yes indee:[d (--------)
```

---

[11] In current dictionary definitions, the prime emphasis is laid on its function as a marker of fact and truth, 'as opposed to *possibly, potentially, theoretically, ideally*; really, in reality' (*Oxford English Dictionary*, 1933), and its sense is also paraphrased as 'strange as it may seem' (*Longman Dictionary of Contemporary English*, 1984). The *OED* states that it is 'not said of the objective reality of the thing asserted, but as the truthfulness of the assertion and its correspondence with the thing; hence added to vouch for statements which seem surprising, incredible, or exaggerated'.

[12] The transcription conventions for CA are discussed in Chapter 2.

(2)   (Clift, 2001:274; H:1:1)
      (L=Lesley, F=Foster. L has rung up F to check that there will be no Sunday school
      that week.)

```
1F   T's a group service'n the evening whi[ch is very suitable=
2L                                         [Yes.
3F   =f'youngsters.
4    (.)
5L   Yes.=I js s-u thought I'd che:ck=
6F   =M[m:.
7L     [I:n case there wz a: misprin:[t. °(Again.)°
8F                                   [Yes no no we're havin:g
9→   ehm: (0.4) w'l I'm away actually b't uh: it's just a group
10   Sundee,
11L  Yes.
```

Here is ample evidence of the 'numerous false starts,[13] deviations from rules,
changes of plan in mid-course' that, if our search is solely for meaning, threatens
to obscure the objects of investigation. However, if instead of treating such data
as 'degraded', we start from the premise that there might be phenomena to be
discovered in them – that we focus on the actions being done in the talk – we can
start, at the very least, by making observations. So, for example, attention to
what are known as repairs and their environment (adjustments or alterations in
the talk directed to problems of hearing, producing or understanding – an issue to
which we return in Chapter 7) reveals that the particle *actually* is implicated in
different ways in the trajectory of the talk. So in (1), 'actually it's g'nna be a
rather busy June' (l. 12) serves to redirect the subsequent trajectory of the talk,
where the lead-up to it, replete with so-called false starts, in the wake of bad
news, had been decidedly delicate. However, in (2), 'w'l I'm away actually'
(l. 9), 'actually' serves to mark the end of a parenthetical insert, after which the
talk resumes its prior topical line 'but uh: it's just a group Sundee'. These two
observations offer only a glimpse of the more extended analysis in Clift (2001),
in which, at one point, we see a single speaker, over a sequence of seventy-eight
lines, producing 'actually' in four different positions in the turn: placements that,
on each occasion, are seen to be wholly systematic, given the actions being
implemented at that given moment (249–51). Thus it is proposed that what
actually does in a stretch of interaction is systematically linked to (a) its position
in a turn, or its component turn-constructional units,[14] and (b) the action
launched by that turn, whether self-repair (as in (1) and (2)), informing, or
topic shift. So in this case the syntactic possibilities exemplified through flexi-
bility of placement are seen to be selected on the basis of interactional exigen-
cies, revealing something of the reflexive relationship between grammatical and
interactional competence.

---

[13] We examine what such 'false starts' can be used to do in Chapter 7.
[14] Turns and turn-constructional units (TCUs) are discussed in Chapter 4.

So, then, when it comes to interaction, it would appear to make as much sense to talk of what *actually* does as what it means. The stable semantic core of contrast and revision (Clift, 2001:286) has its particular sense on each occasion selected by its sequential context, just as the proximity of a boiling kettle selects the type of *drink* being offered.[15] Heritage proposes one possible conceptualisation of this action implication of language in the following terms:

> the appropriate image of a word or a symbol is perhaps that of a large complex organic molecule such as a protein or amino acid existing in three dimensions, in which a variegated profusion of structural configurations and protrusions stand ready to lock into the empirical world of the here and now, stabilizing in the moment (and often for longer) both word and world.   (2011:268)

Examining these moments of stability has much to offer the linguistic attempt to understand the underlying mechanisms of language change and diversity. Heritage, in this connection, invokes the 'philosopher's axe' whose blade and handle have been replaced many times – so it is with language and its components, subject to subtle and imperceptible shifts over generations through myriad interactions between cognitive and communicative constraints. So we would not necessarily recognise the language of our distant ancestors, yet at any moment in time we assume we are speaking the same language: as Sleeth (1982, quoted in Heritage, 2011:67) points out, the Romance languages are the divergent end-products of the gradual, 'imperceptible' change of Latin across time – a temporal shift that has its spatial correlate in the Romance dialect continuum from northern France to southern Portugal, between the far extremes of which there is no mutual comprehension. In this respect, the linguistic work on grammaticalisation mentioned earlier has been critical in tracing how elements of grammar embody a process whereby recurrent formats become sedimented (Bybee, 2001, 2010; Fox, 2007). As Blythe (2013:883), working on Aboriginal languages, notes, repeated selection of particular linguistic structures by reference to interactional and cultural constraints, ultimately leads language down the path to grammaticalisation. Structuration is thus driven by the interactional preference for particular constructions (a topic to which we return in Chapter 5), showing how culture in effect selects for the emergence of structure. How exactly such structures are tailored to the doing of particular actions might suggest that the natural home for such investigation is pragmatic approaches to language use, and in particular those concerned with the distinction between the form and function of utterances: Speech Act Theory, the Gricean theory of utterance interpretation and Relevance Theory.

Pragmatic meaning: three perspectives
(a) Speech Act Theory
Of course, 'doing things with words' has been an object of inquiry in pragmatic approaches to language since Austin's (1962) observations regarding

---

[15] See Clift et al. (2013:210–11) for discussion of an exemplar in the field of colour perception.