# Expressives, Directives and Informatives

## An Analysis of the Evolution of Speech Acts

Author: Hubert Hågemark

Supervisor: Peter Gärdenfors

Master's Thesis (15 credits)

Course code: FTEM01

Department of Philosophy

Lund University

# Table of Contents

# Abstract

Although much research is devoted to the origin of human communication, rather limited efforts have been made to convert these findings into a systematic and overall account of the evolution of speech acts. The thesis contributes to this field by mapping out cognitively relevant features associated with *expressives*, *directives* and *informatives*. Based on empirical data on non-human primate communication, I argue that each speech act class involves diverse cognitive capacities related to acquisition, intentionality, intersubjectivity, involvement, detachment, cooperation and natural meaning. Despite exceptions, the overall results suggest that expressives are the least cognitively demanding, followed by directives and then informatives. This cognitive ladder yields an evolutionary understanding of central speech act classes found in human communication.



*Keywords*: Speech Acts, Animal Communication, Expressives, Directives, Informatives

# 1 Introduction

The influential Oxford philosopher Austin (1962) pointed out that language is a medium for *doing things*. This stands in opposition to the 'traditional' view of language as first and foremost meant to represent and describe the world. Although we indeed use language for those purposes, Austin stressed how we also use it for commanding, ordering, promising, baptizing and expressing emotions – in other words: for doing things. Austin labelled these kinds of activities as *speech acts*. The great influence this work has had in the philosophy of language, pragmatics and semantics, not forgetting fields such as literature studies and experimental psychology (Kissine 2013), is understandable since it casts light on a fundamental aspect of communication. However, despite the extensive study of the evolutionary origin of human communication, rather limited efforts have been made to convert these findings into a systematic and overall story of the evolution of speech acts. This is somewhat surprising given the considerable impact of speech act theory, and especially since such an investigation helps us to better understand the underpinnings of various kinds of communication. The thesis aims to contribute to this field by mapping out *cognitively relevant features* associated with different speech act classes and thereby provide an evolutionary account of them.

The speech act classes under concern are *expressives*, *directives* and *informatives*. Roughly, expressives express the utterer's emotional state; directives encourage the addressee to do something; and informatives encourage the addressee to believe something or attend to some information. Indeed, other speech acts exist, for instance commissives and declarations: in the former the utterer commits to some future action, whereas in the latter she declares something (e.g. 'You're fired') (Searle 1979). Although not exhaustive, the suggested speech act classes form an important starting point for an investigation of the evolution of speech acts. Informatives – as paradigmatic examples of descriptive communication – speak for themselves: they have received considerable attention within the philosophy of language and presumably play a crucial part in our daily cooperative communication. Directives have also been subject to plenty of philosophical discussion. This is understandable given their key role in influencing other people through the means of communication. Expressives, on the other hand, have received less attention within philosophy. However, as noted by Darwin (1890), expressing one's emotional state is a widespread form of communication among humans and other animals. All in all, these three speech act classes emerge as *prima facie* important for the inquiry at hand.

The thesis begins with a theoretical outline of expressives, directives and informatives. This will serve as a point of departure to thereafter structure empirical data on non-human animal (henceforth: animal) communication according to these speech act classes. An analysis of the cognitively relevant features associated with each speech act class will then be provided. This constitutes the main part of the thesis and yields an evolutionary understanding of these forms of communication. Thereafter some concluding remarks will be made.

# 2 Speech Acts

## 2.1 Background

Before going into the various speech act classes, some brief remarks on speech act theory are helpful. The literature on speech act theory is extensive and includes traditions defining speech acts in terms of conventions (e.g. Austin 1962; Searle 1969), Gricean intentions (e.g. Bach & Harnish 1979), common ground (e.g. Portner 2004 & 2007; Stalnaker 1978) and function (e.g. Millikan 1984 & 2005). The most influential work is presumably done by Searle (1979) who develops Austin's account into a more comprehensive and systematized taxonomy of speech acts. Searle suggests

that speech acts largely can be differentiated along their *illocutionary point, sincerity condition* and *direction of fit*.

The illocutionary point is the point or purpose of that type of speech act and forms the most basic unit in his analysis (Searle 1979: 2-3). His colleague Vanderveken understands this as what the utterer necessarily intends to do when she performs that type of speech act (Vanderveken 1985: 183). Although they – especially Searle – do not go very far to substantiate this notion (cf. Siebel 2002), I find it fruitful enough as a tool for differentiating the speech act classes. However, my take on it does not necessitate that the utterer intentionally employs that utterance, since unintentional communication often occurs among animals (see Section 4.2).[1] As a result, if the illocutionary point requires certain intentions, this already from the start excludes plenty of animal communication. The illocutionary point can therefore – vaguely – be understood as what the utterer *necessarily does* when performing a speech act of that type.[2] Importantly, this does not rule out the possibility of the utterance having additional purposes, as long as its primary point remains. I will here be content with these remarks on illocutionary point, partly since the notion will be further clarified in its application, partly since my purpose is not to substantiate this notion.

The sincerity condition concerns which type of psychological state speech acts of a certain class tend to express. Searle holds that different speech acts generally indicate different types of mental states, such as a belief or a desire (Searle 1979: 4-5). By noting the mental state expressed, it is thus possible to gain knowledge of the type of speech act performed (and vice versa).

Speech acts also differ in their direction of fit, i.e. in their relation to the world. Inspired by Anscombe (1957), Searle suggests that some speech acts have an 'utterance-to-world' direction of fit, whereas others have a 'world-to-utterance' direction of fit.[3] Speech acts with an utterance-to-world direction of fit are supposed to make the utterance conform to the world. In other words, the utterance should tell something about the world. Speech acts with a world-to-utterance direction of fit, on the other hand, are meant to make the world conform to the utterance. These utterances thus change the world according to the utterance (Searle 1979: 3-4).

Although these criteria are not exhaustive, they provide a useful basis for further categorizing expressives, directives and informatives.

## 2.2 Expressives

Expressives form a crucial type of speech act among humans and animals. Typical non-linguistic expressives are facial expressions of for instance fear and anger, and vocalizations such as screams and groans (see Section 3 for examples of each speech act class among animals) (Anikin 2020: 13; Bar-On 2013: 317-318). In ordinary language, this includes congratulations, apologies, condolements and greetings, to provide a few examples (Searle 1979: 15). Given that expressives vary from innate natural expressions to highly conventionalized verbal utterances, their illocutionary point is difficult to pinpoint. Focusing on linguistic communication, Searle claims that the purpose of expressives is to express some psychological state of the utterer. For example, when someone apologizes for being late, the utterer expresses her sorrow for this being the case. Searle moreover notes that although the primary purpose is to express some psychological state, this often comes along with providing information about the intentional object of this state, i.e., what the state is about. For example, when someone apologizes for being late, the utterer expresses her

[1] Note that 'utterance' is understood in broad sense incorporating for instance linguistic expressions, gestures and facial expressions (cf. Grice 1989 [1969]: 92).

[2] One might wonder how this take on illocutionary point relates to the various traditions of defining speech acts, for instance in terms of Gricean intentions or function. I will here avoid placing it under any specific tradition but note that it especially bears similarities to a functionalistic take, since it does not require any specific intentions.

[3] Searle himself denotes this in terms of 'word-to-world' instead of 'utterance-to-world' and vice versa (Searle 1979: 3-4). However, in order to also include animal communication, it is suitable to exchange 'word' with 'utterance'.

sorrow and provides information about what she is sorry for, namely that she is late. In Searle's view, the illocutionary point of expressives is thus to express one's psychological state, which often is accompanied by information about its intentional object (Searle 1979: 15; 1983: 16).

An influential voice on expressive communication among animals is Bar-On. Although she is not occupied with speech act taxonomy as such, but expressive communication in general, she presents a similar view. She holds that an expressive both points inward to the psychological state it expresses, and outward toward its intentional object. For instance, an animal's anxious behaviour expresses the experience of fear, together with contextual and or behavioural cues (e.g. gaze) which provide information about its intentional object, such as a nearby predator. In that sense, expressive behaviour both points inward and outward (Bar-On 2015; 2018: 304).

Bar-On and Searle thus unite in that expressives express some psychological state of the utterer, which often is accompanied by information about its intentional object. This additional information is either explicit, as when it is part of the linguistic content of the utterance, or implicit, as when it is revealed by behaviour and gaze directed to some stimuli. However, although expressives often carry or yield additional information, the primary point is still, as Searle emphasizes, to express the utterer's psychological state (Searle 1979: 15).

Bar-On and Searle however fail to distinguish expressives from other speech act categories. Although Searle's examples of expressives often concern the feelings of the utterer, he does not limit the relevant expressed psychological states and appears to include everything from beliefs to desires (Searle 1979: 15; 1983: 173). Similarly, Bar-On understands psychological states as including cognitive, affective and experimental states (Bar-On & Green 2010). The obvious problem with this liberal take on the relevant psychological states is the difficulty to then delimit expressives from other speech act categories. As Searle himself remarks, every type of speech act tends to express some type of psychological state, which he even uses as a criterion for classification, namely the sincerity condition. But since every speech act expresses *some* psychological state, every speech act turns out to be an expressive. This consequence is perhaps unproblematic for Bar-On, who does not put forward an account of speech acts, but peculiar for Searle who does just this.[4]

The remedy for this problem is straightforward: the relevant psychological state of expressives must be the *emotional state* of the utterer. This definition seems to be what Searle somewhat is after since the examples he provides concern the feelings of the utterer (Searle 1979: 15), especially in his later work (e.g. Searle & Vanderveken 1985: 39-40). Bach and Harnish also opt for a definition of expressives (or acknowledgments as they call them) in terms of the utterer's feelings, but with the addon that these must be directed toward the addressee, by for instance expressing regret for having offended the addressee (Bach & Harnish 1979: 51-54). This addon is however problematic since expressives commonly express feelings that are not directed toward the addressee, for instance when expressing anger without being angry at the addressee.

The definition put forward moreover widens the scope of expressives since they do not merely indicate the utterer's feelings but also the utterer's emotional state. Emotional state is here understood broadly including physiological changes, physical sensations (e.g. pain) and feelings. This broader take is appropriate for being able to account for expressives among animals. Thus, the illocutionary point of expressives is to express some emotional state of the utterer, which often is accompanied by information about its intentional object.

After these refinements, the sincerity condition of expressives must be to express some emotional state. Of course, on any such occasion, the emotional state will be of a determinate nature: such as anger, fear or disgust.

Compared to other speech acts, Searle claims that expressives lack a direction of fit. In expressive speech, the utterer neither purports to make the utterance correspond to the world nor make the world conform to the utterance. The utterer simply expresses her emotions, without

[4] Note that this critique presupposes that a speech act only can instantiate one speech act class. Although this should not be taken at face value, it certainly seems to be what Searle thought, and moreover is compatible with the purpose here to group cases of animal communication into various speech act categories.

attempting to either describe or alter the world (Searle 1979: 15). Searle does however not consider that expressives, in some sense, describe the world, namely, the emotional state of the utterer. In contrast to him, I therefore suggest that expressives carry an utterance-to-utterer direction of fit, where the utterance is supposed to conform to the emotional state of the utterer. When the utterer apologizes for being late, she purports to make the utterance match her emotional state.[5] I therefore suggest the following criteria for expressives:

> *Illocutionary point:* express some emotional state of the utterer, which often is accompanied by information about its intentional object.
>
> *Sincerity condition:* some emotional state of the utterer.
>
> *Direction of fit:* utterance-to-utterer.

## 2.3 Directives

Directives have in common that they tend to alter the addressee's prospective behaviour.[6] This includes speech acts such as orders, commands and requests (Jary & Kissine 2014: 2-3 & 13-14). Non-linguistic examples are various gestures, for instance the 'stop hand' gesture, the 'silence' gesture with the index finger, or pointing to make the addressee bring something. In ordinary language this includes primitive utterances such as 'Stop!' or 'Run!' but also full sentences like 'Please sit down' or 'Give me a coffee' (Hamblin 1987: 1-4). These utterances are all important for influencing the addressee's behaviour. Searle for instance holds that a directive 'Counts as an attempt to get H to do A' (Searle 1969: 66). Millikan claims that the 'proper function' of directives is to induce the action expressed, although they do not necessarily succeed in doing it (Millikan 2005: 157-159). According to Bach and Harnish, directives express the utterer's desire for some future action by the addressee, and his intention that the utterance should be taken as a reason for this action (Bach & Harnish 1979: 41). Portner claims that directives add actions to the addressee's 'to-do-list', which contains actions the addressee is expected to perform (Portner 2004: 237; 2007: 352). Although there are crucial differences between these accounts, I take the bottom line to be that directives *encourage* the addressee to do something. Roughly, this means that they make the addressee (significantly) more likely to perform a particular action. Any intentional connotations associated with the term 'encourage' will however be set aside. Compared to Searle, as well as Bach and Harnish, this definition therefore leaves intentionality out of the picture. Moreover, by defining directives in terms of that they merely encourage the addressee to do something, it is not necessary that the addressee always acts accordingly and brings forth the typical perlocutionary effect. This is in line with the above definitions but neither presupposes any specific take on function (see Millikan 2005) nor relies on concepts such as a 'to-do-list' (see Portner 2004).

The sincerity condition associated with directives, i.e. the psychological state they generally express, is that the utterer desires something (Searle 1979: 14). For instance, the utterer desires the addressee to bring a coffee.

Finally, directives have a world-to-utterance direction of fit. They make the addressee act in line with the utterance and thus change the world (Searle 1979: 14). Based on this brief outline, I suggest the following criteria for directives:

---

[5] I gratefully credit Gärdenfors for this suggestion.

[6] The term 'imperative' is often used interchangeably with 'directive'. The former does however not capture the speech act class itself, but only the type of utterance often used to perform a directive (Jary & Kissine 2014: 15; Recanati 2013: 628).

*Illocutionary point:* encourage the addressee to do something.

*Sincerity condition:* that the utterer desires something.

*Direction of fit:* world-to-utterance.

## 2.4 Informatives

Informatives have received considerable attention within the philosophy of language.[7] These are speech acts that in ordinary language commonly are communicated by propositional utterances (Belnap 1990: 1), for instance 'The capital of France is Paris' or 'The king of Sweden is Carl XVI Gustaf'. For the purposes of this thesis, I will presuppose that informatives can occur without the means of declarative sentences. Non-propositional linguistic informatives can be utterances such as when someone says 'Cat!', to make the addressee believe there is a cat on the mat (cf. Bates et al. 1975). Similarly, gestural utterances such as when the utterer instead points to the cat also count as informatives (cf. Brinck 2004: 431). The illocutionary point of informatives is often described in terms of making the addressee believe something. Millikan for instance suggests that the primary function of informatives is to make the addressee create the belief specified by the utterance (Millikan 2005: 157-159). Bach and Harnish suggest that an informative expresses a belief and the intention that the addressee also should undertake this belief (Bach & Harnish 1979: 42). Stalnaker reasons similarly but puts this in terms of common ground. In his view, the point of informatives is to enrich the interlocutors' common ground, i.e. the mutual amount of presuppositions held by the interlocutors (Stalnaker 1978: 323; see also Clark 1996: 38-39 for a similar position). These accounts unite in that the primary purpose of informatives is, roughly, to encourage the addressee to believe something. Bates et al. provide a somewhat weaker definition and understand informatives as speech acts that command the addressee to assume or attend to some piece of information (Bates et al. 1975: 208; Leroy et al. 2009). This definition is more inclusive since, in addition to making the addressee believe something, informatives can also make the addressee attend to some information. Given the purpose to study animal communication, this seems like a desirable result. By merging these accounts, I suggest the primary point of informatives is to encourage the addressee to believe something or attend to some information.[8] [9] Similar to directives, this definition neither necessitates any intentional underpinnings nor that the typical perlocutionary effect always occurs.

The sincerity condition associated with informatives is according to Searle typically that the utterer believes something. This certainly comes in different strengths depending on the subtype of informative, varying from that the utterer knows something (as in assertions) to that the utterer

[7] Some clarifications of terminology are in order. The term informative is intended to capture the type of speech act commonly called 'assertive', 'constative' or 'declarative'. Although 'assertive' often is employed in the field, it is unsuitable for capturing the whole class of informing since it is associated with a 'strong' sense of informing (Kissine 2013: 3). The term 'constative' avoids this problem but is however associated with Austin's distinction (which he later rejected) between utterances that are truth-evaluable, i.e. constatives, and those that are used to perform actions, i.e. performatives (Austin 1962: 3-5; Jary 2010: 7; van Oort 1997; Zhang et al. 2020). 'Declarative' is in turn mainly used to denote the utterance type itself, and not the actual speech act class (Recanati 2013: 629). I will therefore opt for the term 'informative'. This highlights the core feature of this class, namely, to inform, and is vague enough to include various subtypes falling under this umbrella.

[8] Despite the importance of clarification, I will not substantiate the notion 'information' here more than refer to its application in Section 3.3.

[9] By defining informatives in terms of what they encourage the addressee to do (i.e. believe something or attend to some information), they can be understood as a subtype of directives. This solution resembles the definition put forward by Bates et al. (1975: 208).

leans toward something (as in guesses) (Searle 1979: 12-13). I will understand beliefs as wide enough to encompass these variations.

Informatives are the prototypical example of utterances with an utterance-to-world direction of fit. Utterances of this type are primarily used to describe and provide information about the world. In other words, these utterances are supposed to match the world, rather than change it (Searle 1979: 3-4). Thus, informatives typically include the following:

> *Illocutionary point:* encourage the addressee to believe something or attend to some information.
>
> *Sincerity condition:* that the utterer believes something.
>
> *Direction of fit:* utterance-to-world.

### 2.5 Summary of Speech Acts

Although the purpose of the thesis is not to theoretically account for various speech act classes, I believe the hitherto discussion provides a promising basis for exploring animal communication from the perspective of speech act theory. The criteria for the speech act classes are summarised in the following table.

Table 1. Criteria for the Speech Act Classes

| | ***Expressives*** | ***Directives*** | ***Informatives*** |
|---|---|---|---|
| ***Illocutionary point*** | Express some emotional state of the utterer, which often is accompanied by information about its intentional object. | Encourage the addressee to do something. | Encourage the addressee to believe something or attend to some information. |
| ***Sincerity condition*** | Emotional state | Desire | Belief |
| ***Direction of fit*** | Utterance-to-utterer | World-to-utterance | Utterance-to-world |

## 3 Speech Acts among Animals

Before relating the speech act classes to animal communication, it behoves me to make some methodological remarks. Given the purpose of evolutionarily mapping out cognitively relevant features associated with each speech act class, I will primarily consider communication among non-human primates (henceforth: primates). They are evolutionarily close to humans and exhibit diverse and complex forms of communication (Jensvold 2014: 64-65; Leeds & Jensvold 2013: 226). The evolutionary focus will thus be on the phylogenetic development of speech acts and not the ontogeny of humans. The examples provided are moreover selected based on how *typical* they are for each type of speech act. Without getting into controversies about the best take on this notion, it will here be understood as an equilibrium between i) the most common instances of a given speech act and ii) how these examples differentiate one speech act from another (cf. van Rooij & Schulz 2020: 93-94). The same approach will be employed when discussing the cognitively relevant

features underlying these examples. This enables me to provide illustrative examples of each speech act category and their underlying foundations, without getting trapped in necessary and sufficient conditions. However, it also entails that overlapping cases to some extent will be excluded. This constitutes a real problem since a considerable amount of especially primitive communication instantiates core elements of different speech act classes simultaneously. Millikan even labels communicative behaviours like these as 'Pushmi-Pullyu Representations' and believes they are too diverse to be put into any specific category (Millikan 2004: 77; 2005: 166-174). This presumably includes behaviours such as body positioning, gaze or how one animal approaches another, which although communicatively relevant, do not belong to any specific speech act class. As a result, the outline neither purports to provide an exhaustive account of animal communication nor does it claim there are no exceptions. This does however not entail that illustrative examples of the speech act classes do not exist nor that a general story cannot be told. With that said, let us go through instances of expressives.

## 3.1 Expressives among Animals

In Darwin's seminal work (1890), he influentially points out that emotional expressions are widespread among animals. Paradigmatic and prevalent examples are facial expressions and vocalizations (Parr et al. 2005: 716). They are often described as natural expressions, where the link between the emotional experience and the utterance is set by nature (Bar-On 2013: 317-318; 2015).

Facial expressions are common among both humans and animals. Ekman et al. suggest that human facial expressions correspond to emotions such as anger, fear, disgust, sadness, surprise and happiness (Ekman et al. 1972). It is often believed there is an evolutionary continuity between these expressions and the facial expressions among primates (Hess & Thibault 2009: 125-126; Parr et al. 2005: 716). One common facial expression viewed as a precursor to the human smile is the 'play-face' (see Figure 1), which typically is accompanied by play behaviour (Liebal et al. 2014: 86; Parr et al. 2007: 175). An expression that often displays a benign intent is the 'bared-teeth' (Liebal et al. 2014: 85; Parr et al. 2007: 175). Experiences of anxiety and uncertainty are often connected to what is sometimes called an 'alert face' (Parr et al. 2007: 177). Yawning is under certain circumstances an indicator of stress (Paukner & Anderson 2006: 37). Emotions of bonding and physical affinity are associated with the facial expression 'pout' (Parr et al. 2007: 177). These and other facial expressions (e.g. 'whimper') express the emotional state of the utterer (Parr et al. 2007; 2008). They therefore satisfy the illocutionary point of expressives and their sincerity condition. Moreover, since their purpose is to display the emotional state of the utterer, they have an utterance-to-utterer direction of fit.

Vocalizations form another crucial category of expressives (see Figure 2). This includes for instance food, distress and alarm calls. They are commonly understood as having strong emotional underpinnings and mainly expressing the emotional state of the utterer. An influential voice on the matter is Tomasello who claims:

> On the whole, primate vocalizations would seem to be mainly individualistic expressions of emotions, not recipient-directed acts (Tomasello 2008: 19).

Goodall similarly holds that vocal calls almost never appear in the absence of the corresponding emotional state (Goodall 1986: 125). Vocalizations are according to this view deeply connected to the emotional state they express and with limited flexible production. Tomasello tracks this emotional foundation to the urgency often connected to these calls, such as a threat of predators, fights with conspecifics and affiliation with group members (Tomasello 2008: 17-19).

Objections to this picture have however been raised. The crux of the matter is that vocal calls may be tied to emotions, but they nonetheless often convey referential information (cf. McAninch et al. 2009). The classic example of this is when vervet monkeys produce different alarm calls depending on if the predator is a raptor, terrestrial predator or snake (Seyfarth et al. 1980a: 802; 1980b). Seyfarth and Cheney therefore claim that vocal calls are deeply connected to the emotional state they express, but they also provide information about the stimuli that elicit the emotion (although not intentionally so) (Seyfarth & Cheney 2003a: 43 & 51). Recall however that the illocutionary point of expressives is to express some emotional state of the utterer, but that this can be accompanied by information about its intentional object. That vocals call sometimes carry referential information is thus compatible with them being expressives (cf. Kret et al. 2020: 390). Indeed, some calls may be interpreted as 'Pushmi-Pullyu', meaning that they instantiate elements of other speech act classes equally much, and therefore do not count as expressives (Millikan 2005: 166-174). However, this would not imperil the upcoming analysis. First, they would then be exceptions to the general story. As Tomasello remarks, on the whole, vocal calls are primarily expressions of the emotional state of the utterer (Tomasello 2008: 19). Second, since these vocal calls then would be too diverse to belong to any other speech act category, it would not change the analysis of directives or informatives. With that said, vocal calls will primarily be understood as expressions of emotions, even when they carry referential information.

Given that vocal calls tend to express some emotional state, they also fulfil the sincerity condition of expressives. Moreover, since their primary purpose is to display the emotional state of the utterer, they have an utterance-to-utterer direction of fit.

After this brief outline of expressives among animals, we will now continue with directives.

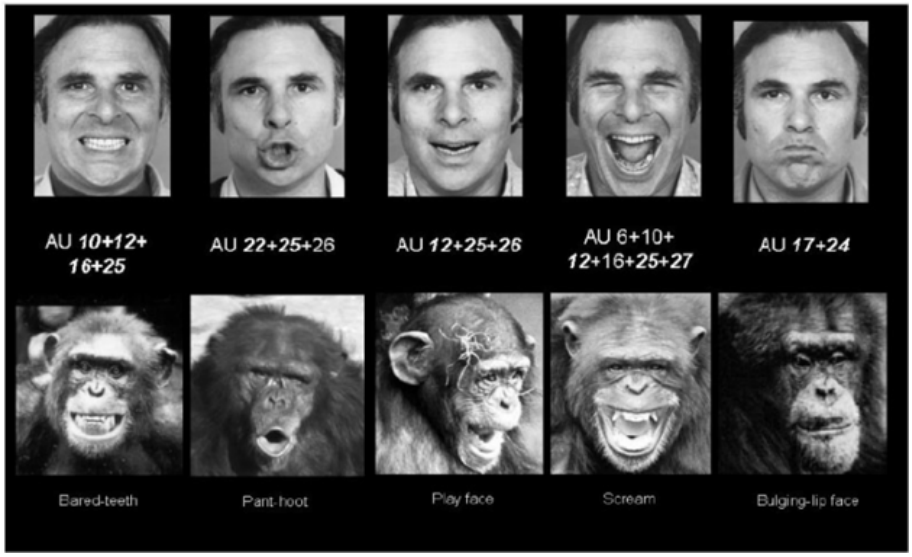


Figure 1. Prototypical chimpanzee facial expressions and homologous facial movements in a human (Ekman et al. 2002).

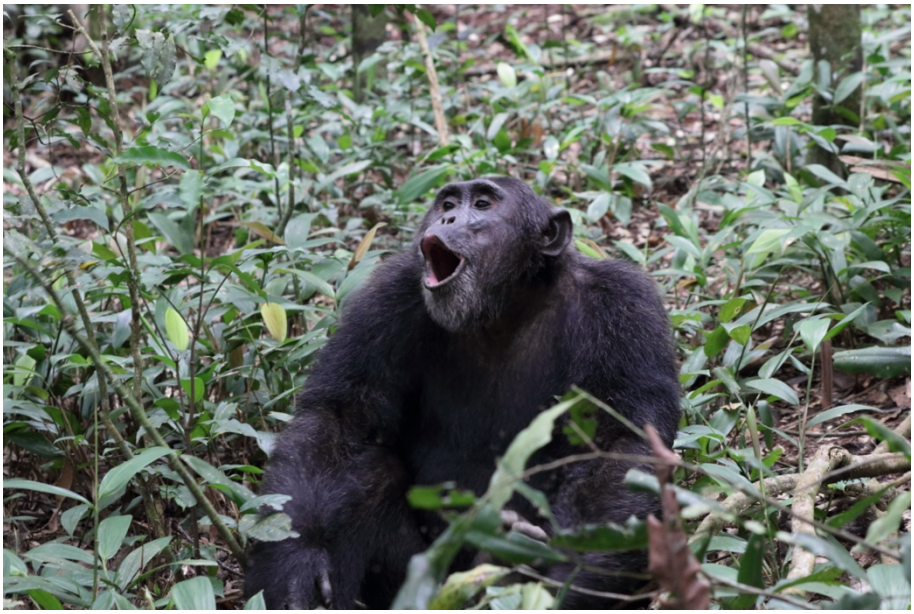

Figure 2. Chimpanzee vocalizing (Taken by Hoult 2018).

## 3.2 Directives among Animals

Directives among animals mostly pertain to various types of gestures. Among the most important are intention-movements, attention-getters, iconic gestures and pointing.

Intention-movements constitute a common type of directive. They consist of gestures where the '[…] individual performs only the first step of a normal behavioural sequence, often in abbreviated form, and this first step is already enough to elicit a response from a recipient […]' (Tomasello 2008: 22). An example of this is the touch-back gesture. In this case, an infant chimpanzee touches the mother's back for the mother to lower it and let the infant climb up (see Figure 3) (see Section 4.1.2 for how this is acquired). Intention-movements are thus used to make the addressee do something.

Although less common, attention-getters are another kind of directive among primates. These gestures consist in that the utterer produces some sort of behaviour that attracts the addressee's attention to a communicative display. A communicative display is usually some physical characteristics of the utterer that naturally carry some meaning or communicative function, for instance facial expressions that display emotions or large horns that deter competitors to approach. The utterer directs the addressee's attention to this display, which in turn communicates something. For instance, young primates who want to play occasionally use attention-getters such as ground-slap, poke-at and throw-stuff, to direct the addressee's attention to their 'play face'. This display in turn communicates that the youngster wants to play. Another example concerns male chimpanzees who attempt to engage in reproductive behaviour. In some regions, the male chimpanzees then occasionally perform a loud behaviour called leaf-clipping, with the purpose to direct the females' attention to their erect penis. In attention-getters, the utterer thus directs the addressee's attention to some display, as an attempt to influence her behaviour (Tomasello 2008: 14 & 27-29).[10]

Directives among primates also involve iconic gestures. They are based on resemblance between the gesture and its reference (Zlatev et al. 2005: 23), be it an object, an action or an event (Russon & Andrews 2011: 627). Pantomime is a subcategory of iconic gestures, where this resemblance is acted out (Gärdenfors 2022: 4; Russon & Andrews 2011: 627). Russon and Andrews

[10] An alternative interpretation of these gestures is that the utterer simply directs the addressee's attention to herself, but not specifically to the communicative display. I will here however go for the standard interpretation provided by Tomasello (2008).

(2011) have documented this behaviour among forest-living orangutans. For example, orangutans occasionally scratch their bodies with a stick and then give the stick to the addressee. They thereby communicate their desire for being scratched. Moreover, if the addressee does not respond appropriately, the pantomime often continues. Other documented cases of pantomime are almost exclusively of a directive kind (Russon & Andrews 2011: 628-629).

Primates brought up in human captivity usually learn to point to make humans do something for them (Leavens & Hopkins 1998: 819). This behaviour especially concerns cases when chimpanzees want food, but also situations when they want the caregiver to for instance open a locked door. Persistence and flexibility often characterize these situations. Chimpanzees are for example able to point at their favourite piece of food (Tomasello 2008: 34-38).

The presented gestures are directives. They encourage the addressee to do something, express the utterer's desire and are supposed to make the world conform to the utterance.

Figure 3. Infant chimpanzee on the mother's back (Taken by Kilby 2010)

## 3.3 Informatives among Animals

It is harder to provide examples of informatives among animals, simply because they seldom occur. Some documented cases can however be interpreted as informatives. They involve language-trained primates that from an early age have been exposed to gestures, human speech and visual symbols (lexigrams) on a keyboard. Apes brought up under this form of human enculturation commonly learn to produce informative utterances by using lexigrams and pointing (see Figure 4) (Lyn et al. 2011: 63-65). To get a picture of the type of informatives involved, I will provide some illustrative examples involving the famous bonobo Kanzi, his half-sister Panbanisha, Panpanzee, a chimpanzee, and the orangutan Chantek. All are language trained and reared in communicatively rich environments.

Informative utterances on keyboard involve comments, showing, labelling, and past and future references. Comments mostly pertain to remarks on ongoing activities without apparent directive intentions. An example is when Panpanzee plays with another ape, goes to the keyboard and presses 'Play', makes eye contact with the human caregiver, and then continues to play. Another example is when Kanzi presses 'Carry' while carrying Panbanisha. Lyn et al. hold that the utterers

here comment on their current behaviour without directive intentions (Lyn et al. 2011: 67-71). On another occasion, Kanzi gives a human caregiver a sweet potato and presses 'Sweet Potato', and then goes on to play again. This is interpreted as a case of showing reference without attempting to influence the addressee's behaviour (Lyn et al. 2011: 67). Labelling is another common form of informative. For instance, while eating strawberries, Panbanisha moves to the keyboard and presses 'Strawberry', and then looks at the human caregiver for confirmation of her labelling (Lyn et al. 2011: 71). Past and future references also occur to some extent. An example is when Panpanzee jumps onto another table and then presses 'Jump', which is understood as a comment on her previous action (Lyn et al. 2011: 67). Although most past references concern events that very recently have occurred, distant references occasionally also occur. For instance, one day Panbanisha sees an injured turtle in the woods, and later in the afternoon points at the symbol 'Turtle'. Lyn et al. interpret the utterer here as commenting on a previous and unusual experience (Lyn et al. 2011: 71).

Informative pointing occasionally also occurs among language-trained primates. Chantek can for instance point toward objects and locations to direct his caregiver's attention. For instance, after being asked about the location of some object he can point correctly at it (Miles 1990: 525; Zlatev et al. 2005: 28). Another case of informative pointing is when Panpanzee points to a jet flying above and then looks at the human caregiver. The utterer is here understood as trying to direct the addressee's attention to some object (Lyn et al. 2011: 72). Pointing and lexigrams are sometimes also combined. For instance, Kanzi has on numerous occasions pointed at some object and then labelled it (Greenfield & Rumbaugh 1991; Lyn et al. 2011: 72).

Although informative utterances are uncommon (Lyn et al. 2011: 69; Tomasello 2008: 38), this outline suggests they occasionally occur.[11] Recall that informatives encourage the addressee to believe something or attend to some information. In these examples, the utterers mainly direct the addressee's attention to some piece of information, such as a current event or the name of an object. However, it is doubtful they encourage the addressee to believe something, except from when they answer questions. These utterances can for that reason be understood as to some extent weak forms of informatives. The sincerity condition of informatives appears satisfied, since these utterances generally express some belief, be it a belief about the name of sweet potatoes or the location of some object. An utterance-to-world direction of fit characterises these utterances given that they are supposed to match the world. For instance, utterances such as 'Carry', 'Play' and 'Jump' are supposed to provide an accurate picture of the current state of affairs. All in all, these utterances fulfil the criteria of informatives.

[11] Note that the corpus Lyn et al. (2011) use primarily is based on spontaneous utterances. If the interaction had been initiated by the human, it would be possible to influence which speech act class the ape would perform, for instance by asking the location of some object. In order to evaluate the speech act classes on equal terms, the discussion is therefore mostly based on self-initiated utterances.

Figure 4. Researcher Sue Savage-Rumbaugh (L), Kanzi (R), and his sister Panbanisha (C) working at the portable 'keyboard' (Taken by Calvin 2006).

# 4 Cognitively Relevant Features

After this outline, I will now continue to the main part and discuss cognitively relevant features associated with the speech act classes. By cognitively relevant features, I mean features that involve or rely on certain cognitive capacities. Their cognitive relevance is often directly apparent, but sometimes more indirect. Recall that the analysis will be based on how typical these features are for each speech act category. It will search for features that are common as well as differentiate one class from another. The suggested features should therefore not be understood as absolute thresholds for each category, but rather as illustrative underpinnings.

## 4.1 Acquisition

The most fundamental question is probably how an utterer acquires the skill to perform a speech act. Let us therefore start from the beginning and explore this.

### 4.1.1 Acquisition: Expressives

Facial expressions and vocalizations among animals are often thought of as unlearned and 'genetically hardwired' (Bar-On 2018: 306; 2021: 10-11; Tomasello 2008: 14-18). This is especially evident in facial expressions. First, within a given species, minor variations of facial expressions exist, and they are often employed in similar social contexts. Second, individuals unable to observe others due to sensory impairments from birth still show typical facial expressions. Third, different primate species often exhibit similar types of facial expressions, despite large variations in social settings. Fourth, the underlying anatomy and musculature are often similar between species (Hess & Thibault 2009: 123; Parr et al. 2005: 716-717; 2008: 216). These circumstances suggest that facial expressions largely are biologically acquired.

The repertoire of vocal calls is to a large extent also genetically based. First, like facial expressions, the variations of vocal calls within a species are often very limited, with minor

individual differences (Fischer & Price 2017: 23; Tomasello & Zuberbühler 2002: 293). Second, social context has a minor impact on the possession of vocal calls. Monkeys raised in absence of conspecifics, either in isolation or among other species, still use the vocal calls typical for their species. Third, efforts to train primates to produce novel vocalizations are generally unsuccessful (Tomasello 2008: 16-17). Fourth, neural evidence suggests limitations in producing novel vocal calls. Primates generally lack equivalent cortical structures that are important for voluntary vocalization among humans (Liebal et al. 2014: 64).

However, despite most vocalisations being genetically acquired, some vocal calls are only found in certain communities and are therefore learned. For instance, the 'raspberry call' is understood as culturally acquired since it is only found among certain populations of gorillas (Lameira 2017: 434). Moreover, even if most call types are innate, some associative learning is occasionally important for learning how to use them appropriately. For instance, infant vervet monkeys are more prone than adults to produce alarm calls in response to harmless animals. They are also unable to discriminate between predatory birds from birds in general. Vervet monkeys therefore learn to use alarm calls specifically in response to predators, although the call type itself is genetically acquired (Seyfarth et al. 1980a: 803; 1980b). The picture that emerges is that expressives are mostly genetically acquired, but occasionally involve some learning.

### 4.1.2 Acquisition: Directives

Directive communication among primates mostly pertains to the use of gestures. They are often learned through social interaction and are not genetically encoded (Tomasello 2008: 20-21).

Intention-movements are commonly acquired through a process called ontogenetic ritualization. This incorporates three stages in the touch-back gesture. First, the infant touches the mother's back and physically attempts to pull it down. Second, after repeated instances of the first stage, the mother starts to anticipate the infant's desire to climb up on her back when she feels the infant's touch. Third, the infant learns to anticipate the mother's anticipation. As a result, the infant uses the touch-back gesture intentionally with the expectation that the mother will lower her back in response to the touch. Other directive gestures commonly involve similar processes of ontogenetic ritualization (Rossano 2013: 166; Tanner & Byrne 1996: 171; Tomasello 2008: 23-25).

Attention-getters are often also learned. First, the utterer must understand that certain displays (e.g. 'play face') carry some communicative meaning. Second, the utterer must note that certain types of behaviours (e.g. ground-slap) attract other's attention, and that they can be used to direct the addressee's attention to a communicative display. These steps often require more demanding learning than ontogenetic ritualization (Tomasello 2008: 27-29).

Tomasello views pointing as a natural extension of attention-getters. The utterer notes that pointing directs the addressee's attention to some referent and then learns to exploit this. It only occurs in interaction with humans since primates are generally not sufficiently cooperative themselves (Tomasello 2008: 37). Brinck however understands directive pointing as a result of ontogenetic ritualization. The utterer first reaches for some nearby object, and then through intermediate steps learns to point at distant objects (Brinck 2004: 432). The acquisition of pointing can therefore be described differently. However, like directives in general, it is not an innate capacity but acquired through learning.

### 4.1.3 Acquisition: Informatives

Informatives occur among enculturated primates. As mentioned in Section 3.3, they are brought up in a stimulating communicative environment with language training and complex interactions with human caregivers on regular basis. They are exposed to various gestures, human speech and lexigrams (Lyn et al. 2010: 361-365; 2011: 63-65). As a result, the acquisition of informatives incorporates extensive learning in a communicatively rich environment.

## 4.2 Intentionality

One often stressed aspect in cognitive science and philosophy of language is the underlying intentions involved in communication. This has its roots in Grice's famous analysis of speaker meaning put in terms of speaker intentions (1989 [1957]). Contemporary scholars have extended his account of speaker meaning to communication in general. This tradition sees communication as a matter of the utterer having certain intentions which the addressee recognizes (e.g. Bach 2012: 53–54; Davis 1992: 240; Haugh & Jaszczolt 2012; Recanati 2012: 135-136; Wilson & Sperber 2006: 611). Given this emphasis on intentions in the literature, it appears relevant to analyse the speech act classes from this perspective. Before going into different types of intentions, it is useful to mention two types of unintentional communicative behaviours. The first is what can be called *involuntary* production, where some communicative behaviour neither is intentional nor possible to suppress. The second is when the utterer can exercise some control over the production but where it is still not based on any intentions to communicate something to the addressee. This can be called *voluntary* production (see Section 4.2.1 for examples) (Green 2007: 28). In contrast to this, we also have intentional communication. Gricean scholars often highlight the *perlocutionary intention*. This is an intention to produce some effect in the addressee, for instance to make the addressee believe or do something (Grice 1989 [1969]; Recanati 1986; Sperber & Wilson 1986: 9-12). Communication is moreover commonly thought of as involving a *communicative intention*: an intention to make the addressee recognize that the utterer has the perlocutionary intention. The communicative intention is accordingly an intention to manifest the perlocutionary intention (Warren & Call 2022: 11). It is this second intention that Grice (1989 [1957]) and several contemporary scholars (e.g. Bach 2012: 52; Heintz & Scott-Phillips 2022: 13; Recanati 1986; Wilson & Sperber 2012: 241) believe separate communication from mere manipulation, since it serves to make communication overt. However, this intention is often criticized for being cognitively too demanding, since the utterer must form the following representations:

> First-order: The utterer intends
> Second-order: That the addressee recognizes
> Third-order: That the utterer intends
> Fourth-order: That the addressee produces the response *p* (believe *p*, do *p*, etc.)
> (cf. Moore 2018a: 5; Thompson 2014: 172; Tomasello 2008: 95).

The communicative intention might therefore be unsuited for dealing with animal communication (Thompson 2014: 172). This motivates a search for other, less demanding, sorts of intentions. One such alternative is what Zlatev et al. call the *communicative sign function*. The utterer then intends an utterance to represent some action, object or event and the addressee to recognize this (Zlatev et al. 2005: 5). On this account, successful interaction requires the interlocutors to assign the utterance the same meaning (Zlatev et al. 2005: 18). Zlatev et al. believe this requires third-order mentality:

> First-order: The utterer knows that *p* means *e*.
> Second-order: The utterer assumes that the addressee knows that *p* means *e*.
> Third-order: The utterer assumes that the addressee knows that the utterer knows that *p* means *e*. (Zlatev et al. 2005: 18)

In my view, the requirement of third-order mentality is unnecessarily restrictive. The importance is that the utterer uses an utterance with a meaning she assumes the addressee knows, which amounts to second-order mentality. Third-order mentality is plausibly beneficial, but the utterer can use a sign communicatively nonetheless. This relaxes the cognitive complexity involved,

especially compared to the communicative intention.[12] The following analysis will relate these three presented intentions, as well as the two types of unintentional communication, to the speech act classes.

### 4.2.1 Intentionality: Expressives

Facial expressions and vocal calls are evolutionarily designed to express the utterer's emotional state. However, they are generally *not* considered as intentional attempts to do this.

The unintentional basis of expressives is most apparent in facial expressions. The emotional state is here considered as the causal basis for the expression (Liebal et al. 2014: 89). Facial expressions manifest for instance the utterer's anger or fear without any intentions to produce these utterances (Green 2007: 88 & 94; Tomasello 2008: 14). Primates who hide their facial expressions with their hands even show that these expressions often are difficult to suppress (which sometimes also happens to humans) (Tomasello 2008: 28). Some control over the production however occurs occasionally. Orangutans can for instance intensify their 'play face' depending on the social context, and therefore exercise some flexibility (Kret et al. 2020: 383). As a result, facial expressions are either involuntary or voluntary, but unintentional.

Although vocalization is deeply tied to the emotional state of the utterer, voluntary control often occurs to some extent. For instance, baboons and diana monkeys can produce or withhold alarm calls depending on the addressee and prior events. Contextual factors are therefore sometimes considered, which indicates flexibility (Seyfarth & Cheney 2003a: 41). This does however often not amount to intentional production. Seyfarth and Cheney view vocal calls as 'inadvertent' since primates generally do not consider how their calls affect the addressee nor do they use them intentionally to inform (Seyfarth & Cheney 2003a: 48; 2003b: 158-159). Vocalization is thus often voluntary but unintentional.

Exceptions to the general story might however exist. Crockford et al. have for instance documented cases where primates appear to produce vocal calls intentionally to inform others (Crockford et al. 2012 & 2015). However, these calls appear uncommon and are presumably best understood as 'Pushmi-Pullyu' and not expressives. They instantiate elements of other speech act classes and it is doubtful their primary purpose is to express emotions (see Section 3.1 for similar reasoning). Recall also that the analysis is based on typical examples of the speech act classes. This means deviant cases to some extent will be excluded. The general picture is that expressives, through facial expressions or vocalizations, are involuntary or voluntary but unintentional.

### 4.2.2 Intentionality: Directives

Directives are commonly viewed as intentional gestures based on a perlocutionary intention. When an infant chimpanzee uses a touch-back gesture, the infant intends to make the mother let her climb on. Similarly, in attention-getters, the utterer intentionally directs the addressee's attention to some communicative display to thereby alter the addressee's behaviour. Pointing and pantomime also involve a perlocutionary intention, and often exhibit persistence if the addressee does not act accordingly (Russon & Andrews 2011: 628-629; Tomasello 2008: 34-38).

The communicative sign function is present in some directives. In attention-getters, the utterer directs the addressee's attention to a communicative display. This requires the utterer to recognize that the display carries some sort of meaning for the addressee, and that this can be exploited for communicative purposes. The same goes for pantomime. When the orangutans scratch their bodies with a stick and then give the stick to the addressee, they must assume that the addressee recognizes the iconic meaning of this utterance. The communicative sign function is however only

[12] For other less cognitively demanding accounts of intentions see for instance Moore (2017 & 2018a) or Gómez (1994).

occasionally employed. Intention-movements – a more widespread type of directive – only involve a perlocutionary intention. These gestures merely require the participants to learn their own communicative role. In the touch-back gesture, the infant only learns to anticipate the mother's response, whereas the mother only learns to anticipate the infant's response. As a consequence, they have different understandings about the meaning of the utterance and their knowledge is not bidirectional (Brinck 2004: 432; Tomasello 2008: 26). Directive pointing works similarly. Typically, apes do not learn the meaning of the pointing gesture itself but only to produce it appropriately. For instance, if a human caregiver points at some place where food is located, they are unable to interpret this successfully. Their knowledge of the gesture is one-sided (Tomasello & Call 2019: 466). Directives, then, involve a perlocutionary intention and occasionally also a communicative sign function.

#### 4.2.3 Intentionality: Informatives

Informatives are clearly intentional. The examples provided in Section 3.3 all include a perlocutionary intention to make the addressee attend to some piece of information, such as a current activity or the location or name of an object. The communicative sign function is also evident in informatives. The enculturated primates intentionally use the symbols on the keyboard to represent some action, object or event, and assume the addressee will recognize this. Their knowledge is bidirectional since they both can produce and comprehend these utterances. Informative pointing also seems to include the communicative sign function. Enculturated apes can not only produce informative pointing but also comprehend pointing much better than other apes (Lyn et al. 2010 & 2011). They therefore understand the meaning of the gesture itself.

Zlatev et al. hold that informatives moreover involve a communicative intention. The reasoning behind this concerns the addressee's ability to recognize the utterer's perlocutionary intention, i.e. the addressee recognizes that the utterer attempts to make the addressee attend to some piece of information (Zlatev et al. 2005: 19-20). However, that the addressee can recognize the utterer's perlocutionary intention does not *per se* entail that the utterer intends the addressee to do this (i.e. the communicative intention). Comprehension is not limited to what the utterer has in mind (cf. Fischer & Price 2017: 29). Although reasonable considering the data, the communicative intention should therefore not be taken at face value. As a result, informatives involve a perlocutionary intention, a communicative sign function, and possibly a communicative intention.

### 4.3 Intersubjectivity

The ability to represent the mental states of others has received considerable attention within philosophy and cognitive science. This is often discussed in terms of having a 'theory of mind', which particularly refers to the capacity to represent the beliefs of others. However, Gärdenfors stresses that this only incorporates an advanced form of mental representation and does not do justice to the diverse spectrum of mental attribution. Gärdenfors instead claims that mental attribution includes representing the emotions of others (empathy); representing the desires of others; representing the attention of others; representing the intentions of others; and representing the beliefs and knowledge of others (theory of mind). These various capacities unite in that they all involve a representation of some form of mental state, an ability Gärdenfors calls *intersubjectivity*. The order of these abilities also represents the cognitive development and evolution of mental attribution, both from a phylogenetic and an ontogenetic perspective (Gärdenfors 2008: 281-286).

Given the cognitive relevance of intersubjectivity, it appears reasonable to explore the speech act classes from this angle. I will begin with the utterer and then move on to the addressee.[13]

### 4.3.1 Intersubjectivity Utterer: Expressives

Based on the hitherto discussion, expressives do not seem to require any intersubjectivity. Although they often affect the mental state of the addressee, they are normally not intentional attempts to do this. Seyfarth and Cheney claim that – compared to human communication – vocalizations among primates typically do not involve the utterer being sensitive to the addressee's beliefs, desires, and so on (Seyfarth & Cheney 2003a: 48). In contrast to gestures, Tomasello similarly suggests that most vocals calls occur without respect to the addressee's attentional state (Tomasello 2008: 18 & 30). Mental attribution is therefore generally not needed in expressives.

### 4.3.2 Intersubjectivity Utterer: Directives

Compared to expressives, directives require the utterer to consider the addressee and her mental states. The utterer must understand the addressee as a causal agent whose behaviour is possible to influence (Brinck 2004: 431; Jary & Kissine 2014: 84-85). Brinck points out the importance of attentional sensitivity in this regard. By recognizing the addressee as an agent, the utterer takes the addressee's body orientation, gaze and behaviour as indicators of her attentional state. This enables the utterer to direct the addressee's attention to the communicative message and to perform the directive act successfully (Brinck 2001; 2004).[14] Tomasello similarly stresses the significance of attentional sensitivity when producing directive gestures. This is especially evident in visual gestures, since they often only occur if the addressee is oriented toward the utterer. For instance, in visual intention-movements, the utterer tends to place herself in front of the addressee before producing the gesture (Tomasello 2008: 31-33). Directives therefore often involve second-order attention, where the utterer notes what the addressee attends to (Zlatev et al. 2005: 17).

An even stronger example of attentional sensitivity is attention-getters. The utterer here not only notes the addressee's attentional state, but also directs it to some communicative display. In addition to attentional sensitivity, this resembles some form of *shared attention*. This type of attention requires the interlocutors to mutually attend to the same object or event, and awareness that the other interlocutor attends to it (see Figure 5) (Zlatev et al. 2005: 17). Although attention-getters are not directed at some external event or object, they resemble this form of attentional awareness. The utterer directs the addressee's attention to some communicative display (e.g. 'play face') and recognizes that the addressee attends to it. The addressee, in turn, presumably notes that the utterer attends to this display. This comes close to shared attention but not all the way since it does not strictly include an external object or event.

Directive pointing however involves shared attention. The utterer and the addressee (i.e. the human caregiver) attend to an external referent, such as a piece of food, and both are probably aware that the other attends to it. Although this only occurs in primate-human communication, it shows that primates are capable of shared attention. Thus, directives often involve attentional sensitivity, which occasionally even amounts to shared attention.

---

[13] Note that intersubjectivity and intentionality partly overlap. However, since the former necessarily involves the mental state of the addressee, whereas the latter concerns what the utterer herself intends, it is reasonable to treat them as two distinct dimensions.

[14] Note that Brinck is primarily occupied with directive pointing (or imperative pointing as she calls it). However, I believe her discussion largely can be applied to directives in general.

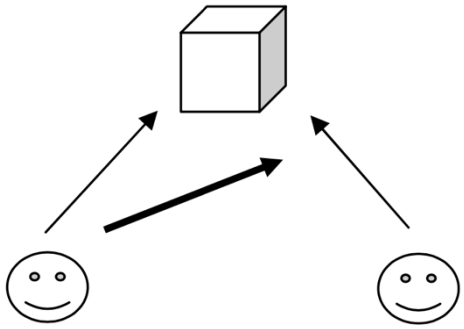

Figure 5. *Shared attention*: Second-order attention: 'I see that you see X' (and vice versa). (From Zlatev et al. 2005).

4.3.3 Intersubjectivity Utterer: Informatives

Informatives involve an even greater amount of intersubjectivity than directives. Given their purpose to make the addressee assume or attend to some piece of information, they certainly incorporate attentional sensitivity. The utterer attempts to make an object or event mutually recognized by the utterer and addressee, which often is thought of as requiring *joint attention*. This form of attention requires the interlocutors to not only note that the other participant attends to some object or event (i.e. shared attention), but also that the other participant recognizes your attention to this. Joint attention therefore entails third-order attention (see Figure 6) (Gärdenfors & Warglien 2013; Zlatev et al. 2005: 17). This type of attention lays the basis for a crucial aspect of informatives, namely, to share reference. The interlocutors mutually attend to some object or event which enables them to exchange information about it (cf. Brinck 2001). Zlatev et al. even claim that the goal of informative pointing is the joint attention itself. The capacity to engage in joint attention is further only documented among enculturated apes (Zlatev et al. 2005: 28 & 34). An illustrative example of an informative utterance with joint attention is the previously mentioned case with the ape Panbanisha who labels strawberries (see Section 3.3). By pressing the symbol for strawberry, the utterer directs the addressee's attention to the strawberries and makes the addressee attentive to her attention to them. Moreover, by then having eye contact with the addressee, the utterer recognizes that the addressee is aware of her attention to them (cf. Moore 2016). Without this joint attention, the utterer would not be able to interpret the addressee's behaviour as a confirmation of her labelling of the strawberries (cf. Brinck 2004: 434). The utterer and addressee establish joint attention, and a similar story can be told for other informatives.

Whether joint attention involves representing others' intentions or beliefs and knowledge states (theory of mind) is however not entirely clear. Brinck argues it does not entail a theory of mind since attentional states are manifested behaviourally, by for instance gaze and body orientation (Brinck 2004: 444). Moreover, because the examples of informatives presented do not have the purpose to induce any beliefs, but merely to draw something to attention, they do not necessitate any theory of mind. The need to represent others' intentions is however difficult to settle. Zlatev et al. claim that joint attention results from a combination of 'I see that you see X' and 'I realize that you want me to look at X' (Zlatev et al. 2005: 19). Although this is a plausible story, it does not show that the utterer had a communicative intention. As already pointed out in Section 4.2.3, the addressee can recognize the utterer's intention without the utterer intending her to do so. Thus, informatives typically involve joint attention, and possibly a communicative intention.

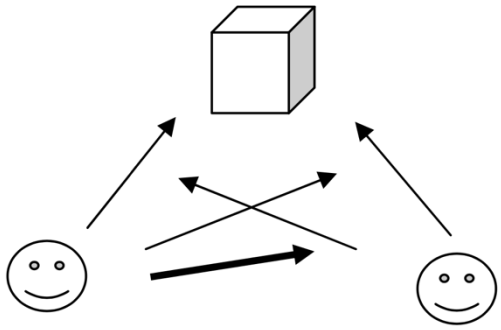

Figure 6. *Joint attention*: Third-order attention: 'I see that you see that I see X' (and vice versa). (From Zlatev et al. 2005).

4.3.4 Intersubjectivity Addressee: Expressives

From the utterer's perspective, expressives do not require any intersubjectivity, as pointed out in Section 4.3.1. However, this tells us little about what is involved in comprehension. Since expressives express some emotional state of the utterer, the addressee must be able to understand that the utterer experiences some emotion, such as being in pain. This does not necessarily involve any cognitive representation, but merely that the addressee generates a somewhat similar emotional state as the utterer (Preston & de Waal 2002: 9). The ability to represent emotions is widely documented among diverse sets of mammal species. This goes hand in hand with that expressives too are widespread among animals (Gärdenfors 2008: 281-286).

Expressives are however produced by an utterer having a certain body position, head-pointing, gaze, etcetera, and sometimes the addressee pays attention to this. By noting these attentional markers, the addressee obtains cues about the stimuli that elicit the emotion, such as the presence of a predator. Expressives are for this reason sometimes accompanied by recognition of the utterer's attentional state (Arnold & Bar-On 2020: 124-125; Bar-On 2018: 303-304). Compared to representation of emotions, this further step is however not necessary. Expressives thus require the addressee to represent emotions, which sometimes is accompanied by attentional sensitivity.

4.3.5 Intersubjectivity Addressee: Directives

Directives are based on that the utterer desires the addressee to do something. Comprehension does therefore presumably often require the addressee to recognize this type of mental state. For instance, in the touch-back gesture, the mother recognizes the infant's desire to climb up on her back. Similarly, in attention-getters where the utterer directs the addressee's attention to a 'play face', the addressee recognizes the utterer's desire to play. This form of intersubjectivity involves more than just representing emotions, since the addressee must note that the utterer has a positive attitude toward some object or event. The addressee must further recognize that the utterer has other preferences than herself (Gärdenfors 2008: 289-290).

Similar to expressives, some attentional sensitivity is also common in directives, such as attention to the utterer's gaze, body position, head tilt, and so on. This is presumably even more important in directives since it helps the addressee to figure out what the utterer wants, and thereby fulfil her perlocutionary intention. As a result, the intersubjectivity involved pertains to the representation of desires and often also attentional states.

### 4.3.6 Intersubjectivity Addressee: Informatives

The comprehension of informatives also requires intersubjectivity. As mentioned in Section 4.3.3, joint attention is an important part of informatives. This requires the interlocutors to not only note that the other participant attends to some object or event, but also that the other participant recognizes your attention to this (see Figure 6). The examples of informatives presented are however based on that the primate produces an utterance directed toward the human. Strictly speaking, these examples do therefore not show the intersubjectivity involved when primates comprehend informatives. However, since they engage in joint attention when they produce informatives, this plausibly characterises their comprehension as well. A common task to investigate comprehension among apes is the object-choice test. The ape is presented with two objects with hidden food in one of them. The human experimenter then points at the object with food and alternates her gaze between the object and the ape. Apes brought up without human language training only perform around chance levels, whereas enculturated apes often successfully conduct the task (Lyn et al. 2010: 362). This communication presumably involves joint attention. Based on the utterer's pointing, the addressee starts to attend to the object and understands that the utterer attends to it. The regular eye contact with the utterer, moreover, makes the addressee aware that the utterer sees that she attends to the object. Joint attention therefore occurs in the comprehension of informatives as well.

As mentioned in Section 4.3.3, Zlatev et al. suggest that joint attention emerges from a combination of 'I see that you see X' and 'I realize that you want me to look at X' (Zlatev et al. 2005: 19). This appears reasonable in light of the object-choice task. The addressee notes that the utterer attends to the object, and, due to the utterer's alternating gaze, understands that the utterer wants her to attend to this object. Comprehension of informatives thus typically involves joint attention and possibly recognition of the perlocutionary intention.

## 4.4 Involvement

Utterances can be structured according to how many participants and referents they involve, namely whether they are monadic, dyadic or triadic. Monadic utterances are self-directed utterances that primarily are about the utterer. Utterances that instead include two individuals interacting with each other, but without any referent, are dyadic. The utterer then directs herself and her communicative message to the addressee. Triadic utterances involve two individuals interacting with each other about a third object, an external referent. In these cases, the utterer links herself, the addressee, and the referent into a 'referential triangle' (Sinha 2003; Włodarczyk 2021: 385-386; Zlatev et al. 2005: 2). In what follows, I show how expressives, directives and informatives can be structured according to their involvement.

### 4.4.1 Involvement: Expressives

Expressives are primarily about the utterer. Through facial expressions and vocal calls, the utterer expresses her own emotional state. These utterances are among primates generally not addressee-directed, since they are not purposefully employed to communicate something to the addressee (Seyfarth & Cheney 2003a: 48; Tomasello 2008: 19). This makes expressives primarily monadic utterances.

The picture is however more complex considering that vocal calls occasionally provide referential information. As mentioned, although they primarily carry information about the emotional state of the utterer, they occasionally provide information about its intentional object (Seyfarth & Cheney 2003a: 43 & 51). However, since the utterer generally does not intend to provide this referential information to the addressee, the utterer does not link herself, the addressee,

and the referent into a 'referential triangle' (Sinha 2003; Zlatev et al. 2005: 2). It is therefore doubtful they can be classified as fully triadic. All in all, it appears plausible to view some of them as *functionally* triadic, since the utterer unintentionally provides referential information. Expressives are therefore primarily monadic utterances, but occasionally also functionally triadic.

#### 4.4.2 Involvement: Directives

In directives, the utterer attempts to influence the behaviour of the addressee, which among animals typically occurs in a dyadic fashion without external referents (Tomasello & Call 2019: 465). This is especially obvious in intention-movements. For instance, in the touch-back gesture, the interaction is solely between the infant and the mother (Tomasello 2008: 23). Similarly, in attention-getters, the utterer attempts to influence the addressee's behaviour by directing the addressee's attention to some communicative display (e.g. 'play face'). Tomasello understands this as dyadic interaction because it does not include an external referent (Tomasello 2008: 27-28). Worth stressing, though, is that attention-getters resemble triadic communication since they involve a communicative display (Tomasello & Call 2019: 466). Directive pointing, however, does involve an external object. Typically, the utterer points to some object (e.g. food) in order for the addressee to bring it for her. This has however been criticized for not being fully referential since the utterer does often not understand the referential function of pointing itself (see Section 4.2.2). It is therefore doubtful whether the utterer links herself, the addressee, and the referent into a 'referential triangle' (Brinck 2004: 433; Tomasello & Call 2019: 465-467). Directive pointing can for this reason be understood as a weak form of triadic communication. As a result, directives are typically of a dyadic kind, although exceptions of weak forms of triadic interactions occur.

#### 4.4.3 Involvement: Informatives

Triadic interaction is most evident in informatives. The utterer draws the addressee's attention to some external object or event and attempts to exchange information about it. Joint attention is often viewed as a basis for full-fledged triadic interaction, since the interlocutors then focus on an object or event for its own sake (Sinha 2003; Zlatev et al. 2005: 19-22). This includes cases of commenting, showing and labelling. A previously mentioned example is when Kanzi gives a sweet potato to the human caregiver and presses 'Sweet Potato' (Lyn et al. 2011: 67). The utterer here shows the addressee an external object and labels it. An example where the reference instead is an event is when Panpanzee presses 'Play' to refer to her previous act of playing (Lyn et al. 2011: 67-71). These interactions count as fully referential: the utterer uses a symbol that refers to some object or event and assumes the addressee will recognize this.

### 4.5 Detachment

The next dimension of evaluation is the degree of detachment (or displacement) involved in the speech act classes. This concerns whether the utterance involves some object or event that is not temporally or spatially present. Detachment is not binary but typically comes in degrees. It can therefore be understood as the degree of context-independence associated with the production of the utterance (Brinck & Gärdenfors: 2003: 488; Gärdenfors 2004: 238; Hockett 1960: 90).

#### 4.5.1 Detachment: Expressives

Expressives primarily point inward to the current emotional state of the utterer. This emotional state is in turn often linked to some ongoing event in the nearby surrounding, such as a nearby

predator (Tomasello 2008: 16-17). Expressives are therefore often tied to the utterer's immediate temporal and spatial environment.

### 4.5.2 Detachment: Directives

Compared to expressives, directives are necessarily future-directed. The utterer wants the addressee to bring forth some desired but not present event. In the touch-back gesture, the infant wants to alter her mother's prospective behaviour. When the utterer uses an attention-getter to direct the addressee's attention to a 'play-face', the utterer wants to start playing with the addressee. Directives are therefore detached in the sense that the event is not present yet. However, the examples presented are based on a present desire for the event to occur directly. Typically, the utterer does not want the event to occur in one hour, tomorrow, or at any subsequent time. The desire for the event is present at the time of uttering. For instance, in the touch-back gesture, the infant wants to climb up on her mother's back straightaway. Similarly, in the attention-getter, the youngster wants to start playing immediately (which is evident given the 'play-face'). Directives among primates are on this account detached in the sense that they are future-directed, but the desire is typically for the event to occur immediately.

### 4.5.3 Detachment: Informatives

Informatives among enculturated primates are typically about some present event or object. The utterer comments, shows or labels something present in their immediate surrounding. However, as mentioned in Section 3.3, past and future references occasionally occur. This typically concerns events that very recently have happened. For instance, when Panpanzee jumps onto another table and then presses 'Jump', she comments on a recent action (Lyn et al. 2011: 67). An example of a more detached reference is when Panbanisha points at the symbol for 'Turtle', although it was several hours ago she saw one. One utterance interpreted as future-oriented is when a human caregiver tells Panpanzee she will be surprised later. During waiting for this surprise, Panpanzee presses the symbol 'Surprise' (Lyn et al. 2011: 70-71). If these interpretations are correct, enculturated apes certainly produce detached informatives. However, they most often refer to events that very recently occurred, and only occasionally to more distant events (Lyn et al. 2011: 70-71). Although they mostly exhibit a weak form of detachment, this is crucial since it constitutes a start for more full-fledged detached interactions.

## 4.6 Levels of Cooperation

Communication is a way to cooperate and coordinate problem-solving (Clark 1996: 62-65). Advanced communication is often viewed as a direct result of an evolutionary pressure of social interaction and having to coordinate actions with others (Melis & Rossano 2022: 2). Cooperation can roughly be defined as agents working together to reach a common goal. On a liberal take, this includes behaviours going from mere instinctive cooperative actions to highly intentional and flexible ones (Brinck & Gärdenfors 2003: 484-485; Gärdenfors 2004: 242). Moreover, a common goal does not necessarily entail that the goal is equally beneficial for each agent or that the agents have similar motives. The importance is that agents work together to achieve something they both find desirable. Given the cooperative value of communication, it appears reasonable to investigate the speech act classes from this perspective.

### 4.6.1 Level of Cooperation: Expressives

Expressives are often associated with urgent needs and situations. The utterer must bond with conspecifics, escape predators, survive fights, and so on. Although evidence suggests that expressives generally benefit the utterer herself (Tomasello 2008: 17), they also help conspecifics to adaptively deal with these urgent kinds of situations (Stephan & Zuberbühler 2021: 167-168). For instance, they generally alert nearby individuals, increase their arousal, and direct their attention to potential threats (Seyfarth & Cheney 2003a: 44-45). Expressives therefore clearly have a cooperative value. However, as previously mentioned, they are generally not employed intentionally, since the utterer does not intend to inform the addressee (Seyfarth & Cheney 2003a: 48). As a result, expressives are unintentionally cooperative.

### 4.6.2 Level of Cooperation: Directives

Directives are based on that the utterer attempts to make the addressee do something for her. Putting it like that, they hardly seem cooperative at all. However, they are part of a cooperative practice where the interlocutors help each other or engage in a mutually beneficial behaviour. This is obvious in the touch-back gesture. Although the infant communicates for purely egoistic reasons, she engages in a practice that mutually benefits the mother (the mother has self-interest in caring for her child). Attention-getters work similarly. When the utterer directs the addressee's attention to a 'play face', she communicates her desire to play. However, the addressee either wants to play herself or has some other motivation for participating; otherwise, she would reject the invitation. The orangutan who pantomimes in order for the addressee to scratch her also participates in a cooperative practice. The interlocutors here presumably engage in an overall reciprocal type of practice where they help each other. Directives are therefore part of an overall cooperative practice important for both interlocutors. However, the utterer's motive is mainly egoistic. As a result, directives can be understood as intentionally cooperative for egoistic reasons.

### 4.6.3 Level of Cooperation: Informatives

Informatives are the prototypical example of cooperative communicative behaviour. The utterer attempts to make the addressee believe something or attend to some piece of information without any direct egotistic motives. The utterer and addressee participate as somewhat equal members and with a common goal to establish joint attention or share information (Brinck 2004: 430; Tomasello 2008: 15 & 38; Tomasello & Call 2019: 467). Take the example mentioned in Section 3.3, when Panpanzee points up to a jet flying in the sky. She tries to direct the addressee's attention to the airplane, without any intention to make the addressee do something for her (Lyn et al. 2011: 67-72). Tomasello even goes so far as to claim that the cooperative underpinning of informatives is what distinguishes human from animal communication. Since animals do not communicate for non-egoistic reasons, they generally do not produce any informatives (Moore 2018b: 333; Tomasello & Call 2019: 467). This is supported by studies on non-enculturated chimpanzees. Generally, they cannot comprehend informative utterances, whereas analogous utterances of a competitive appearance are commonly understood (Lyn et al. 2011: 63-64). Informatives are thus the clearest example of full-fledged cooperative communicative behaviour, where the utterer is intentionally cooperative for non-egoistic reasons.

## 4.7 Natural Meaning

In Grice's seminal paper (1989 [1957]), he distinguishes between what he calls 'non-natural' from 'natural' meaning. Natural meaning is meaning such as 'Those dark clouds mean rain' where 'dark

clouds' somewhat entail that it will rain. This meaning is factive in the sense that the former entails the latter: one cannot consistently say 'Those dark clouds mean rain' and then 'But it will not rain'. Non-natural meaning is instead conventionally tied and includes linguistic meaning. This is meaning such as when someone's hand wave means 'Good to see you!' where the wave sort of conventionally indicates that the person is happy to see you. This meaning is not factive since one can consistently go on and say 'But she isn't happy to see you, she just wants to be nice' (Grice 1989 [1957]: 213-215). In Grice's view, only non-natural meaning is fully communicative and derives from intentional actions, compared to natural meaning which relies on correlations between different types of states of affairs (Borge 2014: 234; Moore 2017; 2018a). The use of non-natural meaning therefore appears cognitively more demanding. However, Grice believes non-natural meaning, through some intermediate stages, ultimately derives from natural meaning (Grice 1989 [1982]: 292-297). Although Grice himself does not provide an evolutionary sketch of such a trajectory, contemporary scholars (e.g. Bar-On & Green 2010; Włodarczyk 2021) have to some extent investigated animal communication through this lens. Unrecognized, however, is how this trajectory partly can be understood by relating it to the various speech act classes under discussion.

### 4.7.1 Natural Meaning: Expressives

Grice himself holds that '[…] things like groans, screeches, and so on […]' normally indicate the emotional state of the utterer and are special cases of natural meaning. These expressions are often biologically anchored and reliable indicators of the emotional states they express (Grice 1989 [1982]: 292). This somewhat fits with the picture of expressives so far painted. Expressives are often deeply tied to the emotional states they express and seldom occur without them. They therefore resemble natural meaning.

Bar-On however emphasizes that even if expressives are reliable indicators of the expressed state, they differ from mere natural meaning. The reasoning behind this concerns the social character of expressives. They carry an intersubjective and overt element where they enable the addressee to recognize the utterer's emotional state and extract valuable information. Compared to non-natural meaning, expressives are however mainly a result of biological processes and lack intentional underpinnings. Bar-On therefore coins the term *social natural meaning* to characterize expressives (Bar-On 2018: 303-305). I believe this nicely captures the natural meaning related to expressives. On the one hand, they are part of natural meaning since they are biologically anchored and are reliable indicators of the expressed state. On the other hand, they have a social and intersubjective foundation. Expressives can consequently be understood as a kind of social natural meaning.

### 4.7.2 Natural Meaning: Directives

Directives incorporate crucial elements of natural meaning but also take a step toward non-natural meaning. They largely result from processes where the interlocutors ritualize or exploit natural meaning. Intention-movements are a clear example of this. After repeated instances of some naturally meaningful interaction, the interlocutors learn to anticipate each other's behaviour by only noting the first stage of this behavioural sequence. The meaning of the gesture is therefore based on a 'preexisting, meaningful social interaction', which the interlocutors learn to anticipate (see Section 4.1.2 for a reminder of how intention-movements are acquired) (Tomasello 2008: 26). These gestures thus result from that natural meaning, in the form of meaningful behaviour, undergoes processes of ritualization.

Attention-getters exhibit similarities. In these cases, the utterer directs the addressee's attention to some display that naturally carries some meaning, such as a 'play-face' or a male chimpanzee's erect penis (see Section 3.2). The meaning of the gesture therefore primarily resides in the natural display itself. The task for the utterer is to grasp, first, the meaning of this display, and second, how

to use it for communicative purposes (Tomasello 2008: 28). Attention-getters are thus based on that the utterer learns to exploit natural meaning.

Pantomime also makes use of natural meaning to some extent. The utterer produces some gesture that resembles a naturally meaningful event and by those means attempts to make the addressee do something. For instance, when the orangutans scratch their bodies and then give the stick to the addressee, they demonstrate the act they want the addressee to perform (Russon & Andrews 2011: 627).

Given that directive pointing can be understood in terms of either attention-getters (Tomasello 2008: 37) or ritualization (Brinck 2004: 432), a similar point can be made about them. Directives are on this account tied to natural meaning, where the interlocutors ritualize or exploit it for communicative purposes.

#### 4.7.3 Natural Meaning: Informatives

Most detached from natural meaning are informatives. In the examples provided, the apes use arbitrary signs on a keyboard to communicate. These are conventional utterances that are acquired through training and are not restricted to any natural meaning. Since these symbols are arbitrary, they do not entail the event or object indicated and are therefore not factive. For instance, an ape can press some symbol without entailing the presence of its reference.

Although the mentioned examples of informatives employ arbitrary signs, this is not necessary. The keyboard could be based on mainly iconic signs where the symbol resembles the reference. Informatives would then exploit natural signs and the gap to natural meaning would shrink. However, the sign would still be conventional since its meaning, although simpler, would have to be learned and be based on mutual knowledge between the interlocutors (cf. Zlatev et al. 2005: 4). It would neither be factive since the symbol would not entail its reference. This makes informatives necessarily conventional.

### 4.8 Summary of Cognitively Relevant Features

Based on the hitherto analysis of the various features associated with the speech act classes, it is possible to summarize the contribution in the following table. Recall that the suggested features should not be understood as absolute thresholds, but as illustrative underpinnings.

Table 2. Cognitively Relevant Features Associated with the Speech Act Classes

| | ***Expressives*** | ***Directives*** | ***Informatives*** |
|---|---|---|---|
| ***Acquisition*** | Innate, occasionally some learning | Learning: ontogenetic ritualization, sometimes more demanding learning | Enculturation: language training and complex interactions with humans |
| ***Intentionality*** | Involuntary or voluntary but unintentional | Perlocutionary intention, sometimes communicative sign function | Perlocutionary intention, communicative sign function, possibly communicative intention |
| ***Intersubjectivity Utterer*** | None required | Representation of attentional state, sometimes shared attention | Joint attention, possibly communicative intention |

| | | | |
|---|---|---|---|
| ***Intersubjectivity Addressee*** | Representation of emotions and sometimes attentional state | Representation of desire and often attentional state | Joint attention and possibly representation of perlocutionary intention |
| ***Involvement*** | Monadic and sometimes functionally triadic | Dyadic and sometimes weak forms of triadic | Triadic |
| ***Detachment*** | None | Future-directed, but a desire for the event to occur immediately | Past and future references, but often about nearby events |
| ***Levels of Cooperation*** | Unintentionally cooperative | Intentionally cooperative for egoistic reasons | Intentionally cooperative for non-egoistic reasons |
| ***Natural Meaning*** | Social natural meaning | Natural meaning is either ritualized or exploited | Conventional acts |

# 5 Concluding Remarks

The empirically based analysis provides a systematic and overall account of the cognitively relevant features underlying expressives, directives and informatives. These speech act classes involve various cognitive capacities related to acquisition, intentionality, intersubjectivity, involvement, detachment, cooperation and natural meaning. Although exceptions and overlaps exist, the results in Table 2 suggest that expressives are the least cognitively demanding, directives hold an intermediate position, whereas informatives are the most complex.[15] Take for example the intersubjectivity involved. In expressives, the utterer does not need to attribute any mental states to the addressee, in directives the utterer needs to consider the addressee's attentional state, and in informatives joint attention is required. Similarly, acquisition goes from mainly innate (expressives), learned through often simple processes (directives), to human enculturation (informatives). The type of involvement required also exhibits this hierarchy, roughly having the following structure: monadic – dyadic – triadic. Cooperation likewise moves from unintentionally cooperative behaviour, to intentionally cooperative but for mainly egoistic reasons, and then peaks in intentionally cooperative for non-egoistic reasons. The same type of cognitive ladder is present in all the remaining dimensions. All in all, the results in Table 2 favour an evolutionary trajectory beginning with expressives, continuing to directives, and then reaching informatives. This evolutionary story is also plausible in light of the prevalence of each speech act class. Expressives are the most common type of speech act among animals, directives occur in various settings but are less common, and informatives mostly occur among primates brought up in human enculturation (cf. Tomasello 2008: 14; Zlatev et al. 2005: 25).

The cognitive capacities associated with the speech act classes however overlap to some extent. For instance, directives commonly involve a perlocutionary intention, but occasionally also the communicative sign function, which typically is present in informatives. Similarly, although directives often are dyadic, they sometimes resemble weak forms of triadic communication. These overlaps are not only compatible with this evolutionary account, but provide a basis for explaining the progression from one speech act class to another.

[15] 'Cognitively demanding' is here roughly understood in terms of how many, complex and higher-order cognitive capacities the various speech act classes require.

Speech act theory casts light on how we do various things when we communicate. Philosophers of language have paid considerable attention to this, but less to the cognitive underpinnings associated with these 'various things we do'. The thesis thus contributes to a philosophically relevant, but, to some extent, unexplored field. However, the evolutionary account put forward is neither exhaustive nor without exceptions. The approach has been to investigate typical examples of the speech act classes and cognitively relevant features associated with them. Since lots of communicative behaviours do not belong to any specific category, they are left outside the story. More research with an empirical angle on speech act theory is therefore needed. This includes studies that have a wider scope on relevant communicative behaviours, as well as research that broadens the data to more animals than primates. Research on the ontogeny of humans is also valuable to explore in relation to speech acts. Another crucial inquiry is to consider other cognitively relevant features. Additional speech acts classes, for instance commissives and declarations, are likewise important to study from an evolutionary point of view. Last but not least, the philosophical implications of these results must be evaluated and discussed.

# 6 Bibliography


Anikin, Andrey (2020), *Human nonverbal vocalizations* [Doctoral Thesis (compilation), Cognitive Science], Department of Philosophy, Lund University.

Anscombe, Elizabeth (1957), *Intention* (2nd ed), Oxford, Basil Blackwell.

Arnold, Kate & Bar-On, Dorit (2020), 'Primate pragmatics, expressive behavior, and the evolution of Language', *Animal Behavior and Cognition* 7(2): 117-130.

Austin, John (1962), *How To Do Things With Words: The William James Lectures delivered at Harvard University in 1955*, edited by Urmson, James, London, Oxford University Press.

Bach, Kent & Harnish, Robert (1979), *Linguistic Communication and Speech Acts*, Cambridge, Mass., MIT Press.

Bach, Kent (2012), 'Saying, meaning, and implicating', In Allan, Keith & Jaszczolt, Kasia (Eds.), *The Cambridge Handbook of Pragmatics*, Cambridge, Cambridge University Press: 47-68.

Bar-On, Dorit & Green, Mitchell (2010), 'Lionspeak: communication, expression, and meaning', In O'Shea, James & Rubenstein, Eric (Eds.), *Self, Language, and World: Problems From Kant, Sellars, and Rosenberg*, Ridgeview Publishing Co.: 89-106.

Bar-On, Dorit (2013), 'Expressive Communication and Continuity Skepticism', *Journal of Philosophy* 110(6): 293-330.

Bar-On, Dorit (2015), 'Expression: Acts, Products, and Meaning', In Gross, Steven, Tebben, Nicholas & Williams, Michael (Eds.), *Meaning Without Representation: Essays on Truth, Expression, Normativity, and Naturalism*, Oxford, Oxford University Press: 180–209.

Bar-On, Dorit (2018), 'Communicative Intentions, Expressive Communication, and Origins of Meaning', In Andrews, Kristin & Beck, Jacob (Eds.), *The Routledge Handbook of Philosophy of Animal Minds*. London, Routledge: 301-312.

Bar-On, Dorit (2021), 'How to do things with nonwords: pragmatics, biosemantics, and origins of language in animal communication', *Biol Philos* 36(50): 1-25.

Bates, Elizabeth, Camaioni, Luigia & Volterra, Virginia (1975), 'The Acquisition of performatives prior to speech', *Merrill-Palmer Quarterly of Behavior and Development* 21(3): 205–226.

Belnap, Nuel (1990), 'Declaratives Are Not Enough', *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 59: 1–30.

Borge, Steffen (2014), 'Horwich on Natural and Non-Natural Meaning', *Acta Anal* 29:229–253.

Brinck, Ingar (2001), 'Attention and the Evolution of Intentional Communication', *Pragmatics & Cognition* 9(2): 255- 272.

Brinck, Ingar & Gärdenfors, Peter (2003), 'Co-operation and Communication in Apes and Humans', *Mind & Language* 18(5): 484–501.

Brinck, Ingar (2004), 'The pragmatics of imperative and declarative pointing', *Cognitive Science Quarterly* 3(4): 429-446.

Calvin, William (2006), Figure 4 - Own work, licensed under CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=49108228

Clark, Herbert (1996), *Using Language*, Cambridge, Cambridge University Press.

Crockford, Catherine, Wittig, Roman, Mundry, Roger & Zuberbühler, Klaus (2012), 'Wild chimpanzees inform ignorant group members of danger', *Current Biology* 22(2): 142-146.

Crockford, Catherine, Wittig, Roman & Zuberbühler, Klaus (2015), 'An intentional vocalization draws others' attention: A playback experiment with wild chimpanzees', *Animal Cognition* 18: 581–591.

Darwin, Charles (1890), *The expression of the emotions in man and animals*, Edited by Darwin, Francis (2nd ed), Cambridge, Cambridge University Press.

Davis, Wayne (1992), 'Speaker Meaning', *Linguistics and Philosophy* 15(3): 223-253.

Ekman, Paul, Friesen, Wallace & Ellsworth, Phoebe (1972), *Emotion in the human face: Guidelines for research and an integration of findings*, Pergamon Press.

Ekman, Paul, Friesen, Wallace & Hager, Joseph (2002), *Facial Action Coding System. Manual and Investigator's Guide*, Salt Lake City, Research Nexus. Figure 2 - Available via license: CC BY-NC 4.0 https://www.researchgate.net/figure/Prototypical-chimpanzee-facial-expressions-and-homologous-facial-movements-in-a-human_fig2_23452365

Fischer, Julia & Price, Tabitha (2017), 'Meaning, intention, and inference in primate vocal communication', *Neurosci Biobehav Rev* 82: 22-31.

Gómez, Juan (1994), 'Mutual awareness in primate communication: A Gricean approach', In Parker, Sue, Mitchell, Robert & Boccia, Maria (Eds.), *Self-awareness in animals and humans: Developmental perspectives*, Cambridge, Cambridge University Press: 61–80.

Goodall, Jane (1986), *The Chimpanzees of Gombe. Patterns of Behavior*, Cambridge, Mass., Harvard University Press.

Greenfield, Patricia & Rumbaugh, Emily (1991), 'Imitation, grammatical development, and the invention of protogrammar by an ape', In Krasnegor, Norman, Rumbaugh, Duane, Studdert-Kennedy, Michael & Schiefelbusch, Richard (Eds.), *Biological and behaviorial determinants of language development*, Hillsdale, Erlbaum: 235-258.

Green, Mitchell (2007), *Self-Expression*, Oxford, Oxford University Press.

Grice, Paul (1989 [1957]), 'Meaning', In *Studies in the Way of Words*, Cambridge, Mass., Harvard University Press: 213–223.

Grice, Paul (1989 [1969]), 'Utterer's Meaning and Intentions', In *Studies in the Way of Words*, Cambridge, Mass., Harvard University Press: 86–116.

Grice, Paul (1989 [1982]), 'Meaning Revisited', In *Studies in the Way of Words*, Cambridge, Mass., Harvard University Press: 283-303.

Gärdenfors, Peter (2004), 'Cooperation and the evolution of symbolic communication', In Oller, Kimbrough & Griebel, Ulrike (Eds.), *The Evolution of Communication Systems*, Cambridge, Mass., MIT Press: 237-256.

Gärdenfors, Peter (2008), 'Evolutionary and developmental aspects of intersubjectivity', In Liljenström, Hans & Århem, Peter (Eds.), *Consciousness Transitions: Phylogenetic, Ontogenetic, and Physiological Aspects*, Elsevier: 281-305.

Gärdenfors, Peter & Warglien, Massimo (2013), 'The development of semantic space for pointing and verbal communication', In Paradis, Carita, Hudson, Jean & Magnusson, Ulf (Eds.), *Conceptual Spaces and the Construal of Spatial Meaning: Empirical Evidence from Human Communication*, Cambridge, Cambridge University Press: 29-42.

Gärdenfors, Peter (2022), 'Teaching as evolutionary precursor to language', *Frontiers in Communication* 7:970069.

Hamblin, Charles (1987), *Imperatives*, Oxford, Basil Blackwell.

Haugh, Michael & Jaszczolt, Kasia (2012), 'Speaker intentions and intentionality', In Allan, Keith & Jaszczolt, Kasia (Eds.), *The Cambridge Handbook of Pragmatics*, Cambridge, Cambridge University Press: 87–112.

Heintz, Christophe & Scott-Phillips, Thom (2022), 'Expression unleashed: The evolutionary & cognitive foundations of human communication: VERSION December 2021', *Behavioral and Brain Sciences*: 1-46.

Hess, Ursula & Thibault, Pascal (2009), 'Darwin and Emotion Expression', *The American psychologist* 64(2): 120-128.

Hockett, Charles (1960), 'The origin of speech', *Scientific American* 203(3): 88–97.

Hoult, Nigel (2018), Figure 2 'Chimpanzee calling', licensed under CC BY 2.0. https://creativecommons.org/licenses/by/2.0/?ref=openverse.

Jary, Mark (2010), *Assertion*, London, Palgrave Macmillan.

Jary, Mark & Kissine, Mikhail (2014), *Imperatives*, Cambridge, Cambridge University Press.

Jensvold, Mary (2014), 'Experimental conversations: Sign language studies with chimpanzees', In Pina, Marco & Gontier, Nathalie (Eds.), *The evolution of social communication in primates: A multidisciplinary approach*, Springer, Cham: 63–82.

Kilby, Eric (2010), Figure 3 'Piggy Back', licensed under CC BY-SA 2.0. https://creativecommons.org/licenses/by-sa/2.0/?ref=openverse

Kissine, Mikhail (2013), *From Utterances to Speech Acts*, Cambridge, Cambridge University Press.

Kret, Mariska, Prochazkova, Eliska, Sterck, Elisabeth & Clay, Zanna (2020), 'Emotional expressions in human and non-human great apes', *Neuroscience and biobehavioral reviews* 115: 378–395.

Lameira, Adriano (2017), 'Bidding evidence for primate vocal learning and the cultural substrates for speech evolution', *Neuroscience and biobehavioral reviews* 83: 429–439.

Leavens, David & Hopkins, William (1998), 'Intentional communication by chimpanzees: a cross-sectional study of the use of referential gestures', *Developmental psychology* 34(5): 813–822.

Leeds, Charles & Jensvold, Mary (2013), 'The communicative functions of five signing chimpanzees', *Pragmatics and Cognition* 21: 224-247.

Leroy, Marie, Mathiot, Emmanuelle & Morgenstern, Aliyah (2009), 'Pointing gestures, vocalizations and gaze: two case studies', In Zlatev, Jordan, Johansson Falck, Marlene, Lundmark, Carita & Andrén, Mats (Eds.), *Studies in Language and Cognition*, Cambridge Scholars Publishing: 261-275.

Liebal, Katja, Waller, Bridget, Burrows, Anne & Slocombe, Katie (2014), *Primate communication: A multimodal approach*, Cambridge, Cambridge University Press.

Lyn, Heidi, Russell, Jamie & Hopkins, William (2010), 'The Impact of Environment on the Comprehension of Declarative Communication in Apes', *Psychological Science* 21(3): 360–365.

Lyn, Heidi, Greenfield, Patricia, Savage-Rumbaugh, Sue, Gillespie-Lynch, Kristen & Hopkins, William (2011), 'Nonhuman primates do declare! A comparison of declarative symbol and gesture use in two children, two bonobos, and a chimpanzee', *Language & Communication* 31: 63-74.

McAninch, Andrew, Goodrich, Grant & Allen, Colin (2009), 'Animal communication and neo-expressivism', In Lurz, Robert (Ed.), *The Philosophy of Animal Minds*, Cambridge, Cambridge University Press: 128-144.

Melis, Alicia & Rossano Federico (2022), 'When and how do non-human great apes communicate to support cooperation?', *Phil. Trans. R. Soc. B* 377.

Miles, Lyn (1990), 'The cognitive foundations for reference in a signing orangutan', In Parker, Sue & Gibson, Kathleen (Eds.), *"Language" and Intelligence in Monkeys and Apes: Comparative Developmental Perspectives*, Cambridge, Cambridge University Press: 511-539.

Millikan, Ruth (1984), *Language, Thought, and Other Biological Categories*, Cambridge, Mass., MIT Press.

Millikan, Ruth (2004), *Varieties of Meaning*, Cambridge, Mass.: MIT Press.

Millikan, Ruth (2005), *Language: A Biological Model*, Oxford, Clarendon Press.

Moore, Richard (2016), 'Meaning and ostension in great ape gestural communication', *Animal Cognition* 19: 223–231.

Moore, Richard (2017), 'Convergent minds: ostension, inference and Grice's third clause', *Interface Focus* 7.

Moore, Richard (2018a), 'Gricean communication, language development, and animal minds', *Philosophy Compass* 13(12).

Moore, Richard (2018b), 'Gricean Communication, Joint Action, and the Evolution of Cooperation', *Topoi* 37: 329–341.

Parr, Lisa, Waller, Bridget & Fugate, Jennifer (2005), 'Emotional communication in primates: implications for neurobiology', *Current opinion in neurobiology* 15(6): 716–720.

Parr, Lisa, Waller, Bridget, Vick, Sarah & Bard, Kim (2007), 'Classifying chimpanzee facial expressions using muscle action', *Emotion* 7: 172–181.

Parr, Lisa, Waller, Bridget & Heintz, Matthew (2008), 'Facial expression categorization by chimpanzees using standardized stimuli', *Emotion* 8(2): 216–231.

Paukner, Annika & Anderson, James (2006), 'Video-induced yawning in stumptail macaques (Macaca arctoides)', *Biology letters* 2: 36–38.

Portner, Paul (2004), 'The Semantics of Imperatives within a Theory of Clause Types', *Semantics and Linguistic Theory* 14: 235-252.

Portner, Paul (2007), 'Imperatives and modals', *Nat Lang Semantics* 15: 351–383.

Preston, Stephanie & de Waal, Frans (2002), 'Empathy: Its ultimate and proximate bases', *Behavioral and brain sciences* 25: 1–71.

Recanati, Francois (1986), 'On Defining Communicative Intentions', *Mind & Language* 1(3): 213-242.

Recanati, Francois (2012), 'Contextualism: Some varieties', In Allan, Keith & Jaszczolt, Kasia (Eds.), *The Cambridge Handbook of Pragmatics*, Cambridge, Cambridge University Press: 135-149.

Recanati, Francois (2013), 'Content, Mood, and Force', *Philosophy Compass* 8: 622-632.

Rossano, Federico (2013), 'Sequence organization and timing of bonobo mother-infant interactions', *Interaction Studies* 14(2): 160–189.

Russon, Anne & Andrews, Kristin (2011), 'Orangutan pantomime: elaborating the message', *Biology letters* 7(4): 627–630.

Searle, John (1969), *Speech Acts: An Essay in the Philosophy of Language*, Cambridge, Cambridge University Press.

Searle, John (1979), *Expression and Meaning: Studies in the Theory of Speech Acts*, Cambridge, Cambridge University Press.

Searle, John (1983), *Intentionality: An Essay in the Philosophy of Mind*, Cambridge, Cambridge University Press.

Searle, John & Vanderveken, Daniel (1985), *Foundations of Illocutionary Logic*. Cambridge, Cambridge University Press.

Seyfarth, Robert, Cheney, Dorothy & Marler, Peter (1980a), 'Monkey Responses to Three Different Alarm Calls: Evidence of Predator Classification and Semantic Communication', *Science* 210(4471): 801–803.

Seyfarth, Robert, Cheney, Dorothy & Marler, Peter (1980b), 'Vervet monkey alarm calls: Semantic communication in a free-ranging primate', *Animal Behaviour* 28(4): 1070–1094.

Seyfarth, Robert & Cheney, Dorothy (2003a), 'Meaning and emotion in animal vocalizations', *Annals of the New York Academy of Sciences* 1000: 32–55.

Seyfarth, Robert & Cheney, Dorothy (2003b), 'Signalers and receivers in animal communication', *Annual review of psychology* 54: 145–173.

Siebel, Mark (2002), 'What Is an Illocutionary Point?', In Grewendorf, Günther & Meggle, Georg (Eds.), *Speech Acts, Mind, and Social Reality. Studies in Linguistics and Philosophy* 79, Dordrecht, Springer: 125–139.

Sinha, Chris (2003), 'The Evolution of Language: From Signals to Symbols to System', In Oller, Kimbrough & Griebel, Ulrike (Eds.), *Evolution of Communication Systems*, Cambridge, Mass, MIT Press: 217-235.

Sperber, Dan & Wilson, Deirdre (1986), *Relevance: Communication and Cognition* (2nd ed), Cambridge, Mass., Blackwell Publishers Inc.

Stalnaker, Robert (1978), 'Assertion', In Cole, Peter (Ed.), *Syntax and Semantics* 9, New York, Academic Press: 315-332.

Stephan, Claudia & Zuberbühler, Klaus (2021), 'Alarm Calling and Kinship', In Shackelford, Todd & Weekes-Shackelford, Viviana (Eds), *Encyclopedia of Evolutionary Psychological Science*, Springer, Cham: 163-171.

Tanner, Joanne & Byrne, Richard (1996), 'Representation of Action Through Iconic Gesture in a Captive Lowland Gorilla', *Current Anthropology* 37: 162–173.

Thompson, Robert (2014), 'Meaning and Mindreading', *Mind and Language* 29(2): 167-200.

Tomasello, Michael & Zuberbühler, Klaus (2002), 'Primate Vocal and Gestural Communication', In Bekoff, Marc, Allen, Colin & Burghardt, Gordon (Eds.), *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition*, Cambridge, Mass., The MIT Press: 293-299.

Tomasello, Michael (2008), *Origins of human communication*, Cambridge, Mass., MIT Press.

Tomasello, Michael & Call, Josep (2019), 'Thirty years of great ape gestures', *Animal Cognition* 22: 461–469.

Vanderveken, Daniel (1985), 'What is an illocutionary force?', In Dascal, Marcelo & Cuyckens, Hubert (Eds.), *Dialogue: An Interdisciplinary Approach*, Amsterdam, John Benjamins Publishing Company: 181-204.

van Oort, Richard. (1997), 'Performative-Constative Revisited: The Genetics of Austin's Theory of Speech Acts', *Anthropoetics* 2(2).

van Rooij, Robert & Schulz, Katrin (2020), 'Generics and typicality: a bounded rationality approach', *Linguistics and Philosophy* 43: 83-117.

Warren, Elizabeth & Call, Josep (2022), 'Inferential Communication: Bridging the Gap Between Intentional and Ostensive Communication in Non-human Primates', *Front. Psychol* 12:718251.

Wilson, Deirdre & Sperber, Dan (2006), 'Relevance Theory', In Horn, Laurence & Ward, Gregory (Eds.), *The Handbook of Pragmatics*, Oxford, Blackwell: 606-632.

Wilson, Deirdre & Sperber, Dan (2012), *Meaning and Relevance*, Cambridge, Cambridge University Press.

Włodarczyk, Mateusz (2021), 'Limitations of Non-Gricean Approaches to the Evolution of Human Communicative Abilities', *Organon F* 28(2): 360-398.

Zhang, Liming, Han, Jie & Li, Jiayi (2020), 'On the Distinction of Constative and Performative Sentences', *Open Journal of Modern Linguistics* 10: 828-833.

Zlatev, Jordan, Persson, Tomas & Gärdenfors, Peter (2005), 'Bodily mimesis as "the missing link" in human cognitive evolution', LUCS: Lund University Cognitive Studies 121.