Demonstratives, joint attention, and the evolution of grammar

Holger Diessel

Correspondence address:
Institut für Anglistik/Amerikanistik
University of Jena
Ernst-Abbe-Platz 8
07743 Jena
Phone: 03641/944530
Email: holger.diessel@uni-jena.de
Demonstratives, joint attention, and the evolution of grammar

Drawing on recent work in developmental and comparative psychology, this paper argues that demonstratives function to coordinate the interlocutors’ joint focus of attention, which is one of the most basic functions of human communication. The communicative importance of demonstratives is reflected in a number of properties that together characterize them as a particular word class: In contrast to other closed-class expressions, demonstratives are universal, they are generally so old that their roots cannot be traced back to other linguistic items, they are among the earliest words that children learn, and they are closely tied to a particular gesture. Moreover, demonstratives play an important role in the internal organization of discourse and the diachronic evolution of grammar, which arguably is also motivated by their communicative function to establish joint attention.

Keywords: demonstratives, joint attention, deixis, pointing, grammaticalization, evolution of grammar
1. Introduction

Linguistic expressions are commonly divided into two basic types, content words and grammatical markers. This is one of the most basic assumptions in linguistics holding across different theoretical approaches. Content words subsume the major word classes, nouns, verbs, and adjectives, whereas grammatical markers comprise a variety of minor word classes such as prepositions, auxiliaries, and determiners. The division between content words and grammatical markers is based on their meanings. Content words have rich semantic interpretations; they denote entities and situations providing the basic content of an utterance. Grammatical markers have schematic (or topographic) meanings; they either ground the meaning of content words (e.g. auxiliaries, determiners) or indicate relationships between chunks of the ongoing discourse (e.g. prepositions, conjunctions). The semantic division between content words and grammatical markers correlates with the size of the various word classes. Content words comprise large and open word classes that can easily be extended by loan words and derivational morphology, whereas grammatical markers comprise small and closed word class that cannot be so easily enlarged by these means.¹

This paper argues that the division between content words and grammatical markers is not sufficient to characterize the basic word classes of human language. Specifically, it shows that there is a special class of linguistic expressions that must be kept separate from both content words and grammatical markers: demonstratives such as English this and that or Japanese sono, kono, and ano.

In the literature demonstratives are commonly classified as grammatical markers functioning as pronouns and determiners; but the current study argues that this does not adequately characterize their function and status in language. Drawing on evidence from linguistic typology, historical linguistics, and language acquisition, I show that demonstratives constitute a unique class of linguistic expressions serving one of the most fundamental functions in language: In their basic use, they serve to coordinate the interlocutors’ joint focus of attention.

Joint attention has been subject to much recent work in developmental psychology, cognitive primatology, and philosophy of the mind (for a review see Dunham and Moore 1995; Krause 1997; Eilan 2005; see also Bruner 1983; Moore and Dunham 1995; Carpenter et al. 1998; Tomasello 1999; Eilan et al. 2005); but although this work has important

¹ There is one class of content words that does not generally comprise a large number of items: In some languages adjectives consist of only a few basic terms for color, age, size, and assessment (Dixon 1982).
implications for the study of language, it has been largely ignored in linguistics. It is the purpose of the current study to show that joint attention plays a foundational role in communication, discourse, and grammar and that demonstratives are commonly used to create (or manipulate) a joint focus of attention.

The paper is divided into three parts. The first part reviews the recent psychological literature on joint attention. It shows that joint attention provides a foundation for the development of communication, social cognition, and language. The second part is concerned with the communicative function of demonstratives. It argues that demonstratives provide a universal linguistic device to coordinate the interlocutors’ joint focus of attention. The communicative importance of demonstratives is reflected in a number of properties that together characterize demonstratives as a particular word class: In contrast to other closed-class expressions, demonstratives are universal, they are generally so old that their roots are not etymologically analyzable, they are among the earliest words that children learn, and they are closely tied to a particular gesture. The third part is concerned with demonstratives in discourse and grammar. It shows that demonstratives play an important role in the internal organization of discourse and the diachronic evolution of grammar, which arguably is motivated by their communicative function to establish joint attention.

2. Joint attention

Joint attention is a complex phenomenon that involves three basic components: the actor, the addressee, and an object of reference. In order to communicate, actor and addressee must share their attention on the same entity or situation. To this end, the actor directs the addressee’s attention to a particular reference object in the surrounding situation; this may involve eye gaze, gesture, or the use of language. If the communicative act is successful, the communicative partners focus their attention on the same referent. In addition, they signal each other through eye gaze or other communicative means that they share the other person’s focus of attention; that is, joint attention requires that the communicative partners recognize they are attending to the same thing (cf. Bruner 1983; Dunham and Moore 1995; Tomasello 1995, 1999; Carpenter et al. 1998; Eilan 2005).

The ability to engage in joint attentional behaviours emerges only gradually during the first year of life (cf. Moore and Dunham 1995; Morissette et al. 1995; Carpenter et al. 1998; Tomasello 1995, 1999; Eilan et al. 2005). Up to the age of 9 to 12 months, children’s interactions are exclusively dyadic; that is, young infants either interact with an adult,
ignoring everything in the surrounding situation, or focus their attention on a particular object, ignoring other people. It is only towards the end of the first year of life that children begin to engage in triadic interactions, in which the child and an adult share attention on a particular referent. The transition from dyadic to triadic interactions is reflected in the emergence of joint attentional behaviours, which can be divided into two basic types: eye gaze and pointing (cf. Eilan 2005).

2.1. The emergence of joint attentional behaviours in child development

In one of the earliest studies on joint attention, Scaife and Bruner (1975) observed that infants as young as 6 months of age begin to follow the gaze of their caregivers (see also Bruner 1983). Subsequent studies confirmed their empirical findings, but emphasized that gaze following at the age of 6 months is only the beginning of a developmental process whereby infants gradually learn to engage in triadic interactions. Butterworth and Jarrett (1991) conducted a series of investigations which led them to conclude that joint visual attention evolves in three successive stages. At the earliest stage, i.e. at 6 months, infants often follow their caregiver’s gaze without identifying the referent; at 12 months they are able to locate the correct target if it is included in the immediate visual domain; and between 12 and 18 months they learn to follow eye gaze to objects even if the object is outside of the immediate visual domain (see also Butterworth 1998; Corkum and Moore 1995; D’Entremont et al. 1997; Carpenter et al. 1998; Franco 2005).

Apart from gaze following, pointing plays an important role in the development of joint attention (cf. Dunham and More 1995; Tomasello 1995; 1999; in press; Butterworth 1998, 2003; Eilan et al. 2005). Deictic pointing is a communicative device that people of all cultures use to establish joint attention (cf. Kita 2003). It usually involves the extended arm and index finger, but there are also other pointing gestures. For instance, in some cultures lip pointing is very common; lip pointing involves the head and the protruded lips pointing in a particular direction (cf. Enfield 2002; Wilkins 2003).

Deictic pointing serves two closely related functions: First, it indicates the location of an object or the direction of movement relative to the location of the pointer; that is, deictic pointing can be interpreted as a guidepost providing spatial orientation. Second, it coordinates the attentional focus of the communicative partners; that is, pointing serves to establish or manipulate a joint focus of attention. In fact, many researchers assume that pointing is “the
quintessential tool for initiating joint attention” (Franco 2005: 139; see also Moore and Dunham 1995; Tomasello 1995; in press; Butterworth 1998, 2003).

The earliest pointing gestures that children produce emerge at around 12 months and typically involve the extension of both arm and index finger towards an object, person, or event (cf. Carpenter et al. 1998; Butterworth 2003; Lizskowski et al. 2004; Franco 2005). Like gaze following, deictic pointing evolves in several successive stages. Before infants begin to use pointing as a communicative device, they often point without intending to focus the adult’s attention on a particular object of reference; the pointing gesture provides orientation in space, but does not function to establish joint attention. This kind of non-communicative pointing has become known as “pointing-for-self”, which precedes “pointing-for-others” (Bates et al. 1976: 217; see also Werner and Kaplan 1963). When children begin to use pointing as a communicative device, i.e. when they begin to point in order to coordinate the shared focus of attention, they gradually learn to combine the deictic pointing gesture with gaze alternation and other joint attentional behaviours (cf. Franco and Butterworth 1996; Franco 2005).

Two types of children’s early pointing gestures are commonly distinguished: proto-imperatives and proto-declaratives (Bates et al. 1976, 1979; see also Bates 1976; Tomasello 1995, 1999, in press; Carpenter et al. 1998; Butterworth 1998, 2003; Camaioni et al. 2004). Proto-imperatives are pointing gestures that resemble reaching gestures produced with the intention to obtain an object (Vygotsky 1926/1962); whereas proto-declaratives are pointing gestures produced with the sole intention to focus the addressee’s attention on a particular object (cf. Bates et al. 1976). Both types of gestures involve the extended arm and index finger, but there is no consensus in the literature as to how and when they evolve. Some researchers have found that infants understand and produce imperative pointing gestures prior to proto-declaratives (cf. Moore and D’Entremont 2001; Camaioni et al. 2004), but other researchers have argued and presented evidence that both types of pointing emerge at around the same time, i.e. at the age of 12 months (cf. Carpenter et al. 1998; Lizskowski et al. 2004).

2.2. Pointing in humans and chimpanzees

Moreover, the two types of pointing seem to involve different social and cognitive skills. Declarative pointing is a unique trait of human communication. While other species may point in one way or another, they do not use declarative pointing gestures (cf. Krause 1997; Tomasello and Camaioni 1997; Butterworth 1998; Tomasello 1999, in press). A number of
studies have shown that chimpanzees and other nonhuman primates may point when they have frequent interactions with humans; but in contrast to humans, chimpanzees use deictic pointing gestures only imperatively; declarative pointing seems to be restricted to humans (but see Leavans and Hopkins 1998 for a different view). Moreover, while nonhuman primates may produce imperative pointing gestures when they have frequent interactions with humans, they are unable to comprehend any type of pointing and do not point for their conspecifics; only in captivity have nonhuman primates been observed to use deictic pointing gestures (for a review see Krause 1997 and Tomasello and Call 1997). There are a few anecdotal reports of wild nonhuman primates pointing for their conspecifics (cf. Veà and Sabater-Pi 1998), but these reports are controversial (cf. Tomasello and Call 1997; Krause 1997; Butterworth 1998). Thus, while chimpanzees and other nonhuman primates are able to learn certain aspects of pointing, they do not point in the same way as humans. – What is it that distinguishes human from nonhuman pointing?

Recent work in developmental and comparative psychology suggests that declarative pointing involves a level of social cognition that nonhuman primates do not achieve even if they have frequent interactions with humans (cf. Tomasello 1995, 1999, in press; Tomasello and Camaioni 1997; Krause 1997; Butterworth 1998; Povinello and Vonk 2003). As pointed out above, declarative pointing gestures are produced with the sole intention to establish joint attention. In order to use such a communicative device, one must be able to recognize that the addressee has mental or intentional states like the self. In other words, joint attention presupposes that the communicative partners understand each other as mental or intentional agents and are able to engage in triadic interactions (Tomasello 1995, 1999, in press; see also Carpenter et al. 1998; Butterworth 1998; Povinelli and Vonk 2003; Franco 2005). According to Tomasello (in press) chimpanzees lack this ability (see also Butterworth 1998, 2003; Povinelli and Vonk 2003). In contrast to humans, they do not understand psychological states in their conspecifics, which is why they do not learn the use of a communicative device that serves to establish joint attention, i.e. a “meeting of minds” (Bruner 1983).

Unlike declarative pointing, imperative pointing does not necessarily involve joint attention. In order to learn imperative pointing, one has to recognize that the use of a deictic pointing gesture induces the addressee to perform a particular action, but one does not have to understand why the addressee reacts in this way; the pointer may conceive of the addressee’s reaction as an automatic response rather than an intentional act. This explains why

2 Like nonhuman primates, children with autism have difficulties in understanding declarative pointing gestures and use pointing primarily in imperative contexts (cf. Baron-Cohen 1995; Leekam 2005).
chimpanzees may learn the production (but not the comprehension) of imperative pointing gestures when they have frequent interactions with humans. What they recognize is that there is some kind of causal link between the pointing gesture and the addressee’s reaction; but they do not understand that the (human) addressee reacts in this way because s/he interprets the pointing gesture as a communicative (i.e. intentional) act.

So then the reason why chimpanzees do not point declaratively is that they do not understand the psychological dimension of human communication (i.e. the “meeting of minds”), which provides an important prerequisite for the ability to engage in joint attentional behaviours. From this perspective, the appearance of joint attention marks a milestone in the ontogenetic development of the child and the phylogenetic development of human beings, providing an important prerequisite for communication and language (Tomasello 1999, in press; see also Butterworth 1998, 2003; Franco 2005).

3. Demonstratives

The earliest behaviours that children use to establish joint attention are pointing and eye gaze; but with the onset of language acquisition, children acquire a new communicative device to manipulate joint attention. In addition to gesture and gaze, language can now be used to create a joint focus of attention. While there are many linguistic means that speakers can use to coordinate a joint attentional focus, there is no other linguistic device that is so closely tied to this function than demonstratives (cf. Clark 1996: 168).

Demonstratives constitute a small class of linguistic expressions that occur in all languages across the world (see below). If we disregard the inflected forms, English has only four demonstratives: this and that and here and there. In the literature, demonstratives are commonly defined as spatial deictics indicating the location of a reference object vis-à-vis the deictic centre. The deictic centre is a conceptual unit that is grounded by the speaker’s location in the speech situation at the time of the utterance (cf. Fillmore 1982, 1997).

While this definition concerns an important aspect of demonstratives, it disregards their communicative function. Demonstratives do not just indicate the location of a referent relative to the deictic centre, they also function to establish (or manipulate) the interlocutors’

---

3 Some studies restrict the notion of demonstrative to this and that and classify here and there as locational adverbs, but the two types of expressions are closely related. Across languages, they share important semantic features and often contain the same deictic roots (cf. Himmelmann 1997; Diessel 1999).
shared focus of attention (cf. Levinson 2004). Thus, I suggest that demonstratives serve two closely related functions:

1. First, they indicate the location of a referent relative to the deictic centre.
2. Second, they serve to coordinate the interlocutors’ joint attentional focus.

In addition, there are certain semantic and syntactic features that are typical of demonstratives. For instance, demonstratives are often deictically contrastive (cf. Engl. *this*/here vs. *that*/there) and serve syntactic functions as pronouns, determiners, and adverbs (cf. Engl. *this*/that vs. *here*/there); but while these features are characteristic of demonstratives, they do not apply to all expressions that are demonstratives according to the above definition (see below). In what follows I show that the communicative function of demonstratives is reflected in a number of properties that together characterize demonstratives as a particular word class.

### 3.1. Demonstratives and deictic pointing

To begin with, demonstratives are commonly accompanied by a deictic pointing gesture (cf. Bühler 1934; Diessel 1999; Enfield 2003; Levinson 2004). The combination of language and gesture is especially characteristic of the ‘exophoric use’ of demonstratives, in which they refer to concrete entities in the surrounding situation (cf. Levinson 2004). In addition to the exophoric use, demonstratives have several other uses, which usually do not involve a deictic pointing gesture (see below); but these other uses are extensions of the exophoric use: They appear later in language acquisition and provide the starting point for the historical development of demonstratives into grammatical markers (Diessel 1999: 109-113; Brugmann 1904: 7-8; Bühler 1934: 390; Levinson 2004).

In the exophoric use, demonstrative serve the same function as a deictic pointing gesture: Both a demonstrative and a deictic pointing gesture indicate the location of an object relative to the deictic centre, i.e. they provide spatial orientation, and both function to focus the addressee’s attention on a particular referent, i.e. they manipulate the interlocutors’ joint attentional focus. While demonstratives and deictic pointing are in principle independent of each other, they are so commonly combined that according to Levinson (2004) the prototypical use of a demonstrative involves a deictic pointing gesture (cf. Brugmann 1904; Bühler 1934; Lyons 1977). In fact, there are languages in which demonstratives “require” an accompanying pointing gesture if they are used with reference to an entity in the physical world (cf. Hellwig 2003: 263; Senft 2004: 62).
The close relationship between demonstratives and deictic pointing provides strong evidence for my hypothesis that demonstratives function to coordinate the interlocutors’ shared attentional focus. In the simplest case, the demonstrative is used to direct the addressee’s attention to a referent that previously was not in the shared attentional focus; in this case, the demonstrative creates a new joint focus of attention. However, demonstratives are also commonly used to direct the addressee’s attention from the current referent to a previously established referent or to differentiate between multiple referents that are already in the shared attentional focus.

In English, demonstratives occur with both functions: *This* and *that* and *here* and *there* may create a new joint focus of attention (e.g. *Look, that’s/there’s Bill*) or may indicate a contrast between two previously established referents (e.g. *Here are two books. This one is mine, and that one is yours*); but in other languages, the two uses are formally distinguished. For instance, Özyürek and Kita (cited in Levinson 2004; see also in Küntay and Özyürek 2002) have shown that Turkish employs the demonstrative *şu* to focus the addressee’s attention on a new referent, while the demonstratives *bu* and *o* are used to indicate a contrast between two referents that are already in the interlocutors’ shared attentional focus. In other words, *şu* creates a new joint focus of attention, whereas *bu* and *o* differentiate between two previously established referents based on their location: *bu* indicates that the referent is closer to the deictic centre than the referent of *o* (cf. Lewis 1967).

### 3.2. The acquisition of demonstratives

Although the deictic features of demonstratives cause difficulties in language acquisition (cf. Clark 1978), children begin to use demonstratives very early. With a few exceptions, the first words that children produce are content words, notably nouns are very common in early child language (cf. Gentner 1982; Clark 2003). One of the few non-content words that children produce during the one-word stage are demonstratives. According to Eve Clark (1978), demonstratives are often among the first ten words that English-speaking children produce, and they are always among the first fifty.

The early appearance of demonstratives is motivated by their communicative function. As pointed out above, demonstratives are commonly accompanied by a deictic pointing gesture, which involves the same communicative constellation and competence as the use of

---

4 English-speaking children need several years to learn the concept of relative distance and the correct interpretation of the deictic centre (cf. Clark 1978; Clark and Sengul 1978; Clark 2003).
language: Both language and pointing occur in triadic interactions involving joint attention and some understanding of the other’s mind. Since the appearance of pointing predates the onset of language, one might hypothesize that children’s early gestures, notably their early pointing gestures, provide the ground for the process of language acquisition (cf. Butterworth 2003).

In order to test this hypothesis, Iverson and Goldin-Meadow (2005) examined the communicative behaviours of ten English-speaking children between the ages of 10 and 24 months. In accordance with the above hypothesis, they found a close correlation between the appearance of gesture and the beginning of language. The earlier children begin to point, the earlier they begin to use language (see also Bates et al. 1979; Desroders et al. 1995; Butterworth and Morisette 1996). Moreover, once children begin to use language, they often combine words with gestures. Two combinatorial strategies appear in the data: Either children use a complementary strategy in which the pointing gesture refers to the same object as the cooccurring word (e.g. pointing to a dog while saying ‘doggy’), or they use a supplementary strategy in which the pointing gesture determines the referent for the predication expressed by the word (e.g. pointing to a doll while saying ‘back’). The former strategy gives rise to a communicative act that corresponds semantically to a one-word utterance, whereas the latter strategy carries the same meaning as a two-word combination (e.g. ‘Dolly back’). Iverson and Goldin-Meadow show that the first appearance of children’s supplementary gestures correlates with the age at which the first two-word utterances appear, providing further evidence for their hypothesis that the use of gesture, notably the use of deictic pointing gestures, “paves the way for language development” (see also Özçalışkan and Goldin-Meadow 2005).

While Iverson and Goldin-Meadow do not discuss the words the children used in their study, their findings support my hypothesis that the early appearance of demonstratives is motivated by their communicative function: Pointing provides the ground for language development, and although pointing is not limited to a particular word class, it seems reasonable to assume that the close relationship of demonstratives to deictic pointing accounts for their early acquisition. As Clark (1978) observed, the earliest demonstratives that children produce are generally accompanied by a deictic pointing gesture (cf. Weissenborn 1988). The combination of demonstrative and deictic pointing creates a powerful tool that allows the child to make reference to any entity in the surrounding situation without knowing the word
for the referent. This does not only explain why demonstratives are among the earliest words children learn but also why they are so frequent in early child language.  

3.3. Universal aspects of demonstratives

The communicative importance of demonstratives is not only reflected in their early acquisition but also in their cross-linguistic distribution. Linguistic typologists have pointed out that most word classes, notably most closed word classes, exist only in a subset of the world’s languages. There are for instance many languages that do not have definite articles, relative pronouns, copulas, auxiliaries, or modal verbs; but there is one notable exception: As has been argued in several recent studies, demonstratives are closed-class expressions that appear to be universal (cf. Himmelmann 1997; Diessel 1999, 2005; Dixon 2003). However, this hypothesis crucially depends on the definition of demonstratives.

In this study, demonstratives are defined as deictic expressions serving two closely related functions: In their basic use, they indicate the location of a referent vis-à-vis the deictic centre and coordinate the interlocutors’ joint attentional focus (see above). Assuming this definition, the currently available data suggest that demonstratives are universal.  

In the literature demonstratives are often defined in terms of their semantic features; notably the expression of distance is often seen as a defining feature of the class (cf. Anderson and Keenan 1985). However, if we define demonstratives in terms of their communicative function (i.e. if we adopt the above definition), we are forced to assume that demonstratives do not generally carry a distance feature (cf. Himmelmann 1996; Diessel 1999, 2005). The Turkish demonstrative şu provides a case in point. Although şu functions to focus the interlocutors’ attention on a particular referent, it does not indicate the relative distance between the referent and the origo (i.e. the deictic centre).

---

5 Examining the speech of four randomly selected 1-to-2 year old English-speaking children from the CHILDES database (cf. MacWhinney 2000), I found 3,857 demonstratives in a corpus of 54,390 words, which means that on the average every 14th word was an instance of this, that, here, or there. Moreover, the demonstrative that was the single most frequent word in the entire corpus (1,794 tokens); the three other demonstratives were less common, but all four demonstratives were among the fifteen most frequent words the children produced (see Table 1 in the appendix).

6 Since we do not have data from all languages it cannot be ruled out that there are languages without demonstratives, but at the moment no such language is known (cf. Himmelmann 1997; Diessel 1999, 2005; Dixon 2003).
Turkish is not unusual in this regard; many languages employ demonstratives that function to direct the addressee’s attention on a referent in the speech situation without indicating its distance to the deictic centre (cf. Diessel 1999, 2005). A well-known example is German, in which demonstrative pronouns and demonstrative determiners do not express a distance feature (cf. Himmelmann 1997; Diessel 1999). If German speakers wish to indicate a contrast between two referents, they simply combine the demonstrative pronouns/determiners with demonstrative adverbs, yielding complex distance-marked forms such *das hier* ‘this one here’ and *das da (drüben)* ‘that one over there’. Similar complex demonstratives have been found in many other languages (cf. Diessel 1999, 2005).

Apart from the expression of distance, the syntactic functions of demonstratives are often included in their definition. In English, demonstratives function as pronouns, determiners, and adverbs; but in some languages we find demonstratives pertaining to other word classes. For instance, Dixon (2003) argues that Dyirbal and Fijian employ demonstratives functioning as verbs, and Diessel (1999) shows that in many languages demonstratives are uninflected particles with no particular syntactic function; notably, the demonstratives in copular clauses often do not fit any of the traditional word classes. For instance, while German has a well-defined class of inflected demonstrative pronouns, functioning as argument of verbs and prepositions, it employs an uninflected particle in copular constructions; the only demonstrative that may appear in German copular clauses is the neuter singular form *das* (cf. 1), which occurs regardless of the gender and number properties of the referent (cf. French *ce*).

(1) **Das** sind meine Schuhe.

DEM.SG.NEUT are my shoe.PL.FEM

‘These/those are my shoes.’

Interestingly, a number of scholars have argued that genuine demonstratives are particles, which developed only later into pronouns, determiners, and other syntactic categories in diachronic change (cf. Brugmann and Delbrück 1911: 311; Bühler 1934: 144; Himmelmann 1997: 21). While it is impossible to verify this view, it is in accordance with recent cross-linguistic findings. There are many languages in which demonstratives cannot be assigned to any of the traditional word classes, and there is evidence that demonstrative pronouns and demonstrative determiners may develop from uninflected particles that acquire their morphosyntactic properties through coalescence with third person pronouns or definite
articles (cf. Diessel 1999: chap 3). Moreover, demonstratives such as English *here* and *there*, which are commonly analyzed as adverbs, are perhaps more appropriately described as particles. Adverbs are linguistic elements that modify the meaning of a verb, but the use of *here* and *there* is not restricted to verbal modification. While *here* and *there* may indicate the location of an activity denoted by a verb (cf. *He lives here*), they are also commonly used in combination with a noun (cf. *this book here*) or preposition (cf. *over there*), which is not consistent with their classification as adverbs.

Thus, if we define demonstratives in terms of their semantic and syntactic features, we would exclude many expressions that are demonstratives according to my definition, and may find that the existence of demonstratives is language-specific (i.e. that some languages lack demonstratives); but if we define demonstratives in terms of their communicative function, we can assume, based on our current knowledge, that demonstratives are universal.

### 3.4. Diachronic aspects of demonstratives

What is more, this universal class of linguistic expressions seems to be much older than other closed-class expressions. In the grammaticalization literature it is commonly assumed that all grammatical markers are diachronic innovations that evolved from content words or from other grammatical markers that previously developed from a content word (e.g. Hopper and Traugott 1993). Since demonstratives are seen as grammatical markers, it is generally assumed that they originate from content words; but despite extensive research there is no evidence for this assumption (cf. Traugott 1982: 245; Hopper 1991: 31; Hopper and Traugott 1993: 129; Diessel 1999: 150-3).

In some languages, demonstratives have been combined with content words to reinforce their pragmatic function. Reinforcement is a mechanism of language change strengthening linguistic expressions that have lost some of their phonetic substance and/or pragmatic force (cf. Lehmann 1995). Since demonstratives are very frequent they are often phonetically and pragmatically reduced. In order to strengthen such a reduced demonstrative it may be combined with other linguistic expressions. Very often, the reinforcing element is another demonstrative (cf. French *celui-ci* vs. *celui-là*); but occasionally a weakened demonstrative is strengthened by a content word such as Latin *ecce* ‘behold’, which reinforced the weakened demonstrative *ille* in Vulgar Latin (*L ille DEM > VL ecce ille > Ofr cest cel > Fr ce*; Harris 1978: 70-78).
While reinforcement is an important mechanism of language change, it does not give rise to a new type of grammatical marker. Even if the reinforcing element becomes part of the demonstrative and the original demonstrative later disappears, the development does not create a new grammatical category. In the end the reinforcing element just continues the function of an old form. If we disregard cases of reinforcement, there is no evidence from any language that demonstratives are historically related to content words; their roots are generally so old that they cannot be traced back to other types of expressions (cf. Brugmann 1904; Traugott 1982; Hopper 1991; Himmelmann 1997; Diessel 1999). Thus, we may assume that demonstratives have emerged very early in the evolution of language so that we simply do not know how they evolved.\footnote{Another class of closed-class expressions that cannot be traced back to content words are interrogatives such as English who and what. Interestingly, Diessel (2003) shows that demonstratives and interrogatives are closely related.}

To summarize, this section has argued that the particular communicative function of demonstratives is reflected in a number of properties that together characterize demonstratives as a particular class: Demonstratives are closely tied to a particular gesture, they are among the earliest words that children learn, they are universal, and their roots are generally so old that they cannot be linked to other linguistic expressions. Concluding this paper I show that demonstratives play an important role in the internal organization of discourse and the diachronic evolution of grammar.

4. Demonstratives in discourse and grammar

4.1. The discourse use of demonstratives

As pointed out above, demonstratives are not only used to focus the interlocutors’ attention on concrete entities in the surrounding situation, they are also commonly used with reference to linguistic elements in discourse (cf. Bühler 1934; Lyons 1977; Himmelmann 1996; Fillmore 1997; Diessel 1999; Wu 2004; Levinson 2004). In the latter use, the deictic centre is shifted from the physical world, i.e. the speaker’s location at the time of the utterance, to a particular point in the unfolding discourse. If we think of discourse as a linear sequence of words and utterances, we may assume that language involves a text-internal origo that is determined, in the string of linguistic elements, by the location of the word that is currently produced (cf. Bühler 1934). Demonstratives that are used with text-internal reference indicate a link
between the linguistic unit in which they are embedded (e.g. NP, PP, S) and the linguistic element to which they refer (e.g. discourse participant, proposition). Two subtypes of the discourse use are commonly distinguished: The anaphoric use in which demonstratives are coreferential with a previous discourse participant (cf. 2), and the discourse deictic use in which demonstratives refer to propositions (cf. 3) (cf. Himmelmann 1996; Fillmore 1997; Diessel 1999; Levinson 2004). 

(2) The Yukon lay a mile wide and hidden under three feet of ice. On top of this ice were as many feet of snow.

(3) Oh, pretty big. Big enough so that the rock doesn't look nearly as tall as it is. The top's bigger than the base. The bluff is sort of worn away for several hundred feet up. That's one reason it's so hard to climb.

The demonstrative in (2) is an anaphoric demonstrative that indicates a coreference relationship between the current NP and a referent in the previous sentence, and the demonstrative in (3) is a discourse deictic demonstrative that refers to the preceding propositions. Anaphoric and discourse deictic demonstratives serve different discourse pragmatic functions: Anaphoric demonstratives keep track of prior discourse participants, whereas discourse deictic demonstratives establish links between chunks of the ongoing discourse (cf. Himmelmann 1996).

Since anaphoric and discourse deictic demonstratives do not refer to concrete entities, they are usually not accompanied by a pointing gesture; however, they involve the same psychological mechanisms as demonstratives that speakers use with text-external reference. In both uses, demonstratives focus the interlocutors’ attention on a particular referent. In the exophoric use they focus the interlocutors’ attention on concrete entities in the physical world, and in the discourse use they focus their attention on linguistic elements in the surrounding context. In other words, in both uses demonstratives function to create a joint focus of attention. Joint attention is thus not only important to coordinate the interlocutors’ attentional focus in the speech situation, it also plays an important role in the internal organization of discourse.

---

8 In addition to these uses, there are other more specialized uses in which demonstratives do not have an immediate referent in the surrounding discourse (or speech situation). For instance, many languages employ demonstratives to indicate that speaker and addressee are familiar with an entity or situation due to shared experience (cf. “… and then he did that little raised eyebrow thing …”) (cf. Himmelmann 1996).
4.2. The development of demonstratives into grammatical markers

Moreover, when anaphoric and discourse deictic demonstratives are routinely used to express a particular relationship between two linguistic units, they often lose their deictic force and develop into grammatical markers. Across languages, demonstratives provide a common historical source for some of the most frequent grammatical expressions such as definite articles, relative and third person pronouns, complementizers, conjunctions, copulas, and focus markers. While some of these expressions have lost their referential force, the developments are motivated by the fact that demonstratives function to establish joint attention. In the remainder of this paper I briefly discuss some of the grammaticalization processes whereby demonstratives develop into grammatical expressions (for a more detailed discussion see Diessel 1999: chap 6).

One of the most frequent grammatical markers that commonly develops from a demonstrative is the definite article. Definite articles occur in many languages across the world and are almost always derived from a demonstrative. For instance, the English definite article the developed from the demonstrative þēs in Old English (Traugott 1992; see also Krámský 1972; Ultan 1978; Greenberg 1978, 1991; Harris 1980; Cyr 1993; Laury 1997; Himmelmann 1997, 1998; Diessel 1999, Heine and Kuteva 2002). The development of the definite article originates from an appositional construction in which a demonstrative pronoun (or particle) is only loosely adjoined to a cooccurring noun. In this construction, the demonstrative directs the addressee’s attention to a linguistic or non-linguistic element, and the noun provides semantic information that can help the addressee to identify the referent. If this construction is routinely used, it may develop into a hierarchically organized NP in which the demonstrative turns into a determiner of the head noun (cf. Himmelmann 1997; Diessel 1999). Parallel to the syntactic development, the demonstrative may assume a new pragmatic function. The pragmatic development originates from an anaphoric demonstrative referring to a previous discourse participant. Like anaphoric demonstratives, definite articles indicate a coreference relationship between the current NP and a previous discourse referent; but they occur in somewhat different contexts. In languages in which demonstratives and definite articles are formally distinguished, definite articles signal the continuation of a currently activated discourse participant, whereas anaphoric demonstratives indicate a topic shift or a contrast between two previously established discourse referents (cf. Himmelmann 1996; Comrie 1998; Diessel 1999). Moreover, while both anaphoric demonstratives and definite
articles are used to track a previous discourse participant, they have other uses in which they do not overlap (cf. Hawkins 1978; Lyons 1999; Himmelmann 1997). For instance, definite articles do not occur with text-external reference (unless the referent has been mentioned in the previous discourse), suggesting that when demonstratives develop into definite articles they are limited to text-internal reference. Parallel to this development, definite articles acquire new uses in which the referential function is often backgrounded. For instance, definite articles are commonly used in the “anaphoric-associative context” (Hawkins 1978), in which the definite NP is not coreferential but semantically associated with a previous discourse participant (e.g. *There was a house. The door was open.*). Another innovation is the use with singletons (e.g. *the sun*), which in many languages require a definite article even if the referent has not been mentioned before. In the end, the demonstrative turns into a marker of pure definiteness signalling that the addressee should be able to identify the referent (cf. 4).

(4)  deictic > anaphoric > definite

Like definite articles, third person pronouns commonly develop from anaphoric demonstrative pronouns (cf. Bhat 2005). For instance, the English third person pronouns *he* and *it* can be traced back to demonstratives (cf. Traugott 1992). The development of third person pronouns is similar to the development of definite articles. In both cases, the development is motivated by the functional overlap between source and target. Like anaphoric demonstrative pronouns, third person pronouns continue a previously established discourse participant, but tend to appear in somewhat different contexts: Anaphoric demonstrative pronouns are commonly used to differentiate between two discourse participants or to indicate a topic shift, whereas third person pronouns are commonly used to continue the current discourse topic (Comrie 1998; Diessel 1999: chap 5). The division between the two categories is fluid and there are many languages in which demonstrative pronouns and third person pronouns have the same form (cf. Bhat 2005). The categorical split is commonly introduced through the complementary distribution of stress accent: Anaphoric demonstrative pronouns tend to be stressed, whereas third person pronouns are usually unstressed, which may lead to phonological reduction and hence to the emergence of a new form. If the development continues, third person pronouns may evolve into clitics, which may turn into agreement markers, before they eventually disappear (cf. 5) (cf. Diessel 1999: 119-120).

---

9 The development of she is unclear (cf. Traugott 1992).
Another frequent grammaticalization pathway leads from anaphoric demonstrative pronouns to relative pronouns. For instance, the German relative pronoun *der/die/das* developed from an anaphorically used demonstrative pronoun (cf. Behagel 1923-1932, III: 766). Both types of pronouns function to continue a previous NP; but due to diachronic change the use of relative pronouns is syntactically more constrained. Demonstrative pronouns turn into relative pronouns when the sentence in which they occur is reanalyzed as a subordinate clause, which usually involves the development of a more rigid word order: Relative pronouns tend to be more restricted in their distribution than demonstrative pronouns in that they must occupy the first position in the clause continuing the immediately preceding NP. Demonstratives are more flexible in this regard; they may occur in any position in the clause and the antecedent may be separated from the pronoun by intervening (clause-level) constituents (cf. 6). If the grammaticalization process continues the relative pronoun may lose its referential function and develop into a pure formal marker of the relative clause.

English *that* provides a case in point. Originally, *that* was an inflected relative pronoun, but in Modern English it has lost its function as a pronoun: relative *that* is non-referential, uninflected, and can be omitted.

Anaphoric demonstrative pronouns also provide a common historical source for copulas. Since the development does not occur in European languages, it is little known in linguistics despite the fact that this process appears to be quite common (cf. Li and Thompson 1977; Schuh 1983; Gildea 1993; Devitt 1994; Diessel 1999). For instance, Li and Thompson (1977) have shown that the Chinese copula *shì* developed from a demonstrative pronoun that occurred in a topicalization construction consisting of a fronted (i.e. topicalized) NP and a nonverbal clause in which the demonstrative resumes the topicalized element. This construction developed into a copular clause in which the topicalized NP has been reanalyzed as subject and the demonstrative as copula (cf. 7).

---

10 Two alternative explanations are reviewed in Diessel (1999: 121-3).

11 In subject relative clauses the relative marker is obligatory, but in all other types of relative clauses it can be omitted (e.g. *The book that amused Peter* vs. *The book (that) Peter found amusing*).
Closely related to the development of copulas is the development of focus markers (cf. Diessel 1999: chap 6). Many languages employ cleft-constructions to present a referent in focus position. Cleft constructions are complex sentences consisting of two clauses: One clause presents the focused element, and the other clause expresses an assertion. In English, the focused element is presented in a copular clause, but many languages do not employ copulas in this construction, using instead a nonverbal clause consisting of the NP in focus position and a pronominal element that is based on a demonstrative. When this construction is routinely used to present an NP in focus position, making it available for the assertion expressed in the subsequent clause, the pronominal element may lose its deictic function and turn into a focus marker (cf. 8) (cf. Diessel 1999; Heine and Kuteva 2002).

Finally, demonstratives provide a common historical source for complementizers and sentence connectives. In this case, the development originates from a discourse deictic demonstrative referring to a proposition. For instance, the complementizers of North and West Germanic languages developed from a demonstrative that functioned to anticipate the occurrence of a subsequent clause (cf. Lookwood 1968; Hopper and Traugott 1993; Harris and Campbell 1995; Diessel 1999). The source construction can be exemplified by the following sentences from Modern English: ‘I can tell you this: ‘I am not going to this party.’” The example consists of two clauses that are combined by the demonstrative this referring from the end of the first clause to the proposition of the second clause. In the Old Germanic languages, constructions of this type developed into complex sentences in which the demonstrative was reanalyzed as a complementizer that became associated with the second clause, i.e. the complement clause, where it occurs in the Modern Germanic languages (cf. 9).  

Very often, the source construction included a copy of the cataphoric demonstrative at the beginning of the second clause. While such ‘correlative constructions’ are somewhat different from the example discussed in this section, the grammaticalization process is similar (cf. Hopper and Traugott 1993; Diessel 1999).
Like complementizers, subordinate conjunctions and sentence connections may develop from a discourse deictic demonstrative (cf. Diessel 1999; Heine and Kuteva 2002). Very often, the development originates from constructions in which the demonstrative is embedded in an adpositional phrase creating a thematic link between two sentences. For instance, Modern German has a particular class of pronominal adverbs that developed from an old oblique form of the demonstrative das and an adposition (e.g. *damit* ‘DEM.with’, *dabei* ‘DEM.by’; cf. English *therefore, hereby*). The composite forms function syntactically as adverbs that are commonly used to express a thematic link between two clauses (e.g. *Wir haben den Zug verpasst; darum sind wir zu spät.* ‘We missed the train; therefore we are too late.’). The construction is only weakly grammaticalized: The preposition is still governed by the verb and the demonstrative has preserved its referential force (cf. 10).

\[ (10) \quad [\ldots]_S \ [\text{[DEM} P]_{PP\ldots} ]_S \quad \rightarrow \quad [\ldots]_S \ [\text{[DEM.P]ADV\ldots} ]_S \]

In addition to the grammatical categories discussed in this section, there is a wide variety of other grammatical markers that commonly evolve from a demonstrative: temporal adverbs (cf. Anderson and Keenan 1985), directional preverbs (cf. Diessel 1999), linking articles (Himmelmann 1997), nominal and verbal number markers (cf. Frajzyngier 1999), possessive pronouns (cf. Diessel 1999), determinatives (cf. Diessel 1999), expletives (cf. Diessel 1999), and topic markers (cf. Vries 1995).

It must be emphasized that the grammaticalization of demonstratives involves some of the most frequent and most important grammatical phenomena that we find in languages across the world: definiteness marking, reference tracking, relativization, complementation, clause combining, and the formation of copular clauses—all of these phenomena are cross-linguistically very frequent and commonly derived from constructions including a demonstrative. Thus, although demonstratives constitute small classes of linguistic expressions, their contribution to the evolution of grammar is significant; grammar would look very different without this source.

5. Conclusion

To summarize, this paper has argued that demonstratives constitute a special class of linguistic expressions that serve a particular communicative function and play an important role in the internal organization of discourse and the diachronic evolution of grammar. In the
literature demonstratives are commonly analyzed as grammatical markers, but this analysis is not appropriate to characterize their function and status in language. Drawing on evidence from developmental and comparative psychology, I have argued that demonstratives function to coordinate the interlocutors’ joint focus of attention. The ability to engage in joint attentional behaviours provides an important prerequisite for communication and language. In order to communicate speaker and addressee must focus their attention on the same referent and must recognize the triadic constellation of the situation. In other words, demonstratives serve one of the most basic communicative functions in language. This is reflected a number of properties that together characterize demonstratives as a particular class:

1. Demonstratives are closely tied to the gestural communicative system. There is no other class of linguistic expressions that is so closely associated with a particular type of gesture than demonstratives.
2. Demonstratives emerge very early in language acquisition. They are often the first non-content words that children learn based on the prior use of deictic gestures.
3. Demonstratives are universal. In contrast to other closed-class expressions, they occur in all languages across the world.
4. Demonstratives are generally so old that their roots are not etymologically analyzable. In contrast to other closed-class expressions they cannot be traced back to content words.

Further, we have seen that the communicative function of demonstratives is commonly extended from the physical world to the universe of discourse. Demonstratives are not only used with reference to concrete entities in the surrounding situation, they may also refer to linguistic elements in the ongoing discourse. While the discourse use is more abstract than the exophoric use, it involves the same psychological mechanism; in both uses demonstratives function to create a joint focus of attention. In the exophoric use they focus the addressee’s attention on a concrete entity in the physical world, and in the discourse use they direct the addressee’s attention on a linguistic element in discourse.

Finally, we have seen that when demonstratives are routinely used with text-internal reference they often develop into grammatical markers. Across languages, demonstratives provide a common historical source for definite articles, third person pronouns, relative pronouns, complementizers, sentence connectives, copulas, and a wide variety of other grammatical markers. While some of these expressions have lost their deictic force, their
developments are motivated by the communicative function of demonstratives to establish joint attention.

All this suggests that demonstratives constitute a special class of linguistic expressions. They serve one of the most basic functions in language, i.e. they establish/manipulate joint attention, which is not only important to coordinate the interlocutors’ communicative interactions, but also plays a key role in the internal organization of discourse and the diachronic evolution of grammar. Since joint attention provides an important prerequisite for the development of communication and language, it seems reasonable to assume that demonstratives have emerged very early in the evolution of language. They are part of the basic vocabulary of every language and must be kept separate from all other linguistic expressions.
### Appendix

Table 1. Number of tokens of the most frequent words in the speech of four 1-to-2-year old English-speaking children from the CHILDES database

<table>
<thead>
<tr>
<th></th>
<th>Eve</th>
<th>Naomi</th>
<th>Nina</th>
<th>Peter</th>
<th>Total</th>
<th>%mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td><strong>that</strong></td>
<td>860</td>
<td>327</td>
<td>241</td>
<td>366</td>
<td>1794</td>
</tr>
<tr>
<td>2.</td>
<td><strong>it</strong></td>
<td>481</td>
<td>488</td>
<td>142</td>
<td>303</td>
<td>1414</td>
</tr>
<tr>
<td>3.</td>
<td><strong>a</strong></td>
<td>581</td>
<td>97</td>
<td>234</td>
<td>349</td>
<td>1261</td>
</tr>
<tr>
<td>4.</td>
<td><strong>there</strong></td>
<td>299</td>
<td>175</td>
<td>52</td>
<td>500</td>
<td>1026</td>
</tr>
<tr>
<td>5.</td>
<td><strong>the</strong></td>
<td>340</td>
<td>145</td>
<td>341</td>
<td>74</td>
<td>900</td>
</tr>
<tr>
<td>6.</td>
<td><strong>my</strong></td>
<td>348</td>
<td>61</td>
<td>314</td>
<td>161</td>
<td>884</td>
</tr>
<tr>
<td>7.</td>
<td><strong>what</strong></td>
<td>146</td>
<td>511</td>
<td>10</td>
<td>162</td>
<td>829</td>
</tr>
<tr>
<td>8.</td>
<td><strong>no</strong></td>
<td>353</td>
<td>138</td>
<td>117</td>
<td>115</td>
<td>723</td>
</tr>
<tr>
<td>9.</td>
<td><strong>mommy</strong></td>
<td>283</td>
<td>187</td>
<td>148</td>
<td>29</td>
<td>647</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>13.</td>
<td><strong>this</strong></td>
<td>41</td>
<td>406</td>
<td>52</td>
<td>97</td>
<td>596</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>15.</td>
<td><strong>here</strong></td>
<td>67</td>
<td>31</td>
<td>247</td>
<td>96</td>
<td>441</td>
</tr>
<tr>
<td>Total</td>
<td>20.512</td>
<td>13.072</td>
<td>8.551</td>
<td>12.255</td>
<td>54.390</td>
<td>100</td>
</tr>
</tbody>
</table>
References


Call, J. and Tomasello, M. 2005. What chimpanzees know about seeing, revisited: An explanation of the third kind. In Eilan et al. (eds.), 45-64.


