12 Deixis

Christoph Rühlemann and Matthew Brook O'Donnell

12.1 Introduction

Talk, taken out of context, has little meaning. For those who participate in it, talk reveals its full and specific meanings only against the "background" of the context in which it occurs (Goodwin and Duranti 1992: 3): the said requires, for its interpretation and its analysis, the frame of the unsaid (Goffman 1974). This unsaid can take many different guises: what speakers said earlier (previous text), what they know of the world (world knowledge), of their culture (cultural knowledge) and of each other (social knowledge), how they interact non-verbally (non-verbal behavior), and what they intend to communicate (illocutionary act), to name only a few. Another central component of the contextual frame is the spatio-temporal situation in which an utterance is made. While most other contextual components remain in the background (that is, in the unsaid), the situational context does come to the fore in the said. It becomes apparent in deixis. Deixis comprises "those features of language which refer directly to the personal, temporal or locational characteristics of a situation within which an utterance takes place" (Crystal 2003: 127). Deixis is linguistic evidence of how what is said is grounded in the context of the situation in which it is said. It provides an interface linking language and situational context (Hanks 1992: 48; see also Lyons 1977: 636, Hanks 2011: 315).

In Section 12.2 we characterize deixis in some more detail. Section 12.3 reports on a case study on what is called "introductory *this*," a usage of *this* which is specific to conversational narrative. While Section 12.2 uses examples collected from a wide range of corpus sources, the case study in Section 12.3 is based on, and illustrated by, data from the Narrative Corpus, a corpus of conversational narratives (see Rühlemann and O'Donnell 2012). Section 12.4, finally, presents some conclusions and points for future development.

12.2 Characteristics of deixis

12.2.1 Deixis and reference

According to C. S. Peirce's widely used distinction, signs can be either iconic, symbolic, or indexical. A sign which is an index indicates a referent via an "existential relationship" (see Levinson 2004: 102) that holds between the referent and the sign, as for instance smoke indicates fire. Deictic words, or deictics, form a subclass of indexical language use in that they are "ready-made indexical expressions" (Levinson 2004: 102). The reference they accomplish is exophoric; that is, deictics such as *you*, *tonight*, and *at my house* are merely "signposts" (Bühler 1990 [1934]: 93) pointing out a relation of some sort (personal, temporal, locational, etc.) that holds between the speaker and the situation (Crystal 2003: 127; Roberts 2004: 199). Being signposts, the meanings of these forms are "altogether 'abstract'" (Sacks 1992: 520): they "contribute to individuated referring without in any way describing their objects" (Hanks 2011: 316). Deictic reference can thus only be resolved in relation to the speech situation: depending on the situation, *you* can refer to any of the current speaker's interlocutor(s), the reference *tonight* "points" to the night of the day of speaking, and *at my house* singles out the speaker's domicile. Given this context-dependence (Levinson 1983, Marmaridou 2000, Crystal 2003), it is not surprising that, for children, "deixis is as confusing as a hall of mirrors: my 'I' is your 'you', my 'this' is your 'that', my 'here' is your 'there', and so forth" (Levinson 2004: 99) and that, for example, the demonstratives are used correctly only from the age of four (Tanz 1980: 145). Also, deictic reference varies not only between speakers but, potentially, even within speakers. As speakers move through space and time, so do their deictic references to the same objects of reference (see Hanks 1992, 2009). For example, "[w]hen a speaker construes his home as '(over) there' in one utterance and 'right here' in the next, he individuates the same place but under two different perspectives" (Hanks 2009: 21–22).

To complicate matters, the context-dependent deictic system interacts in intricate ways with the context-independent descriptive system in which it is embedded. That is, "[t]o use Peirce's terminology, we have an intersection of the indexical plane into the symbolic one – it's a folding back of the primitive existential indexical relation into symbolic reference" (Levinson 2004: 99). Due to this embedding in the symbolic system, typical deictic terms can lose their deictic character while others, typically symbolic terms, can gain deictic status. As regards inherently deictic words used non-deictically, consider the demonstratives *this* and *that* in the vague expression *this and that* (see O'Keeffe et al. 2011: 38), and *there*, for example, in *there you go*, where the idiomatic usage effaces any locational reference.

(1) And she says come on, my, you girl I tell you **this and that**. (BNC-C: KCV 5484)

(2) That's the spoon right, **there you go** look plastic spoon (BNC-C: KBW 15924)

As regards the possibility of "indexicalizing" normally non-indexical expressions, there seems to be no limit, for "just about any referring expression can be used deictically" (Levinson 2004: 101) provided they are used with an immediate view to situational circumstances.[1] Consider Levinson's (2004: 101) examples:

(3) **The funny noise** is our antiquated dishwasher (said pointing chin to the kitchen)

(4) **What a great picture!** (said looking at a picture)

12.2.2 Deixis and anaphora

Treating typically non-deictic material as if it were deictic can also be observed in the use of third-person pronouns, which are mostly used anaphorically. Traditionally, anaphors have been seen as "continu[ing] a pre-established reference in text by pointing back to a specific antecedent, usually an NP [noun phrase]" (Schwarz-Friesel and Consten 2011: 355). In (5), for example, the referent of the pronouns *his* and *he* can be determined due to the aforementioned proper noun *Fowler*. While "deixis and anaphoricity are prototypically incompatible phenomena" (Marmaridou 2000: 73), a borderline case is what might be called "out of the blue" pronominal usage between familiars, where third-person pronouns refer "to referents specified, not in the same conversation, but in a conversation held maybe one or two or even more days earlier" (Rühlemann 2007: 62), and where the temporal distance between the conversations would normally undercut the possibility of backward reference. In the invented example in (6), for instance, the speaker has just discovered the passport of her son, who has departed for a tour abroad, lying on the kitchen table. In the utterance in (6), made to her husband upon entering the kitchen, the references *he* and *his* can only be said to be referring backward if the spouses have been talking about their son before entering the kitchen.

(5) Fowler regained **his** composure and confidence when **he** broke O'Connor's serve in the fourth game of the third set . . . (BNC: K4T 10647)

(6) **He**'s forgotten **his** passport! (invented example)

[1] The possibility of using non-indexical referring expressions indexically shows why "indexicality exceeds the bounds of ready-made indexical expressions, i.e. deictics with in-built contextual parameters" (Levinson 2004: 103).

12.2.3 Deixis and gesture

Deictic forms typically include expressions "that are semantically insufficient to achieve reference without contextual support" (Levinson 2004: 103). The key to "loading" these forms with the contextual support they require is "the direction of the addressee's *attention* to some feature of the spatio-temporal physical context" (Levinson 2004: 102, added emphasis; see also Hanks 2011: 316). To secure this attention, deictics often "co-articulate with gestures" (Hanks 2009: 12), including pointing gestures but also reduced forms such as directed gaze or a nod (Levinson 2004: 102). But deictics need not always be used with gestural support ("gestural usage"), as in (7), where the speaker's utterance is likely accompanied by a size-demonstrating gesture (see Fillmore 1997: 64). Deictics can also be used symbolically ("symbolic usage"), as in (8), where no gesture is needed for the listener to identify the country being referred to (see Levinson 1983: 65–66, Fillmore 1997: 62–63).

(7) And so the flowers are like daisies **this big**. (COCA: NBC_Today)

(8) **This country** is going bankrupt. (COCA: Fox_Hannity)

12.2.4 The deictic origo

Given their context-dependence, deictic words "presuppose an origo relative to which they are computed" (Hanks 1992: 50). Traditionally, the "origo," "deictic center" or "zero-point," has been associated with the current speaker. The origo can metaphorically be understood as the "viewpoint" from which the speaker relates to the dimensions of person, time, and place (Lyons 1977: 638) and which is defined as "NOW in time, HERE in place and FIRST SINGULAR in person" (Mindt 2000: 16; see also Bühler 1990 [1934]: 117). Deictic reference thus has three parts: "a referential focus (the object), an origo from which it is picked out (the indexical ground), and the relation between the two" (Hanks 2009: 11). Where several deictics are used in an utterance they normally all match the speaker's origo or viewpoint, thus forming a coherent "deictic context" (Lyons 1977: 638), or "deictic system" (Levinson 1983: 68). The coherence in the deictic context is best illustrated where speakers use related deictics from the same deictic category. Consider (9), in which the speakers are talking about birthday presents. Adrian is addressing his interlocutor – hence the vocative *Karen* and the second-person pronoun *you* – and referring to the birthday present belonging to that referent – hence the second-person possessive pronoun *yours*. In Karen's response, attention is focused on the same referents (viz. Karen and her present), but given the switch from Adrian's to Karen's origo, the deictics are now the first-person pronouns *I* and *mine*.

(9) ADRIAN: **Karen** why aren't **you** opening **yours**?
KAREN: **I**'ve opened **mine**.

(BNC: KD0 4627–4628)

Typically, then, the organization of deixis is egocentric (Levinson 1983: 63; Marmaridou 2000: 70) and the origo changes when speaker change occurs (Levinson 1983). However, the "ego," or viewpoint, is by no means always the current speaker's "ego," or viewpoint. Speakers can slip into alter egos and organize deixis from that non-ego viewpoint: what Lyons (1977: 579) refers to as "deictic projection" and Fillmore as a "switching of deictic centers" (Fillmore 1997: 122). Deictic projection occurs in multiple forms. Perhaps the most immediately obvious form is found in "child-directed speech," viz. the use by adults of referring expressions conforming to a child's viewpoint. Consider (9), where Christopher and Tim are children, and Dorothy and Andrew their parents. In the arrowed utterance, Dorothy refers to her husband twice as *daddy*, the kinship term appropriate to the children's viewpoint; the projection is upheld even as she addresses her husband directly (cf. the tag *aren't you daddy?*). Clancy (2010: 128), working on spoken corpora of Irish English, argues that in parent–child interaction, the "family deictic centre is constructed around the children".

(10)
CHRISTOPHER: Cottage pie?
TIM: All of us?
CHRISTOPHER: Yeah.
ANDREW: Not me cos I'll be at work. But the rest of you.
CHRISTOPHER: I love cottage pie.
→ DOROTHY: But **daddy**'s having lunch at work, **aren't you daddy**?
ANDREW: I think so.

(BNC-C: KBW 10834-10841)

Further, complexities in terms of origo arise in the use of motion deictic verbs such as *come* and *go*, which are said to encode "movement to the speaker" (*come*) and, respectively, "movement from the speaker" (*go*) (Levinson 1983: 83–84; Fillmore 1997: 77–102). However, in certain contexts, *come* is standardly used to mean "movement to the addressee." Consider:

(11) I'm coming

It would be highly non-idiomatic to replace (11) by "I'm going to where you are." In saying *I'm coming*, which "implicitly encodes the role of the addressee" (Marmaridou 2000: 93), the speaker assumes the addressee's perspective.

Another classic diversion from speaker-egocentricity is found in constructed dialogue, also referred to as direct speech, as in (10), where the reporting

speaker, despite his using the words *I* and *you*, is not referring to himself and, respectively, his interlocutor (see Yule 1996: 15). Instead, the reference is "dislocated" (Hanks 2011: 318) or "projected" from the reporting speaker to the reported speaker and his interlocutor as well as from the present speech event to the prior speech event in which the quoted utterance *I'll give you a minute* (purportedly) occurred.

(12) he said "**I**'ll give you one minute" (NC: KC1-N1)

Rühlemann (2007: 68) argues that one of the reasons that in most spoken corpora *I* and *you* are high up among the most frequent words is their use in constructed dialogue in conversation not to refer to the present speaker and their interlocutor but to a distal speaker in an anterior situation and their equally removed addressee.

Further, use of the first-person plural pronoun *we* standardly vacillates with regard to the underlying origo. As noted by Biber et al. (1999: 329), *we* usually "refers to the speaker/writer and the addressee (inclusive *we*), or to the speaker/writer and some other person or persons associated with him/her (exclusive *we*)." That is, while the use of exclusive *we* is computed relative to an origo which is not shared by the addressee, inclusive *we* projects an interpersonally shared origo. The exclusive use is illustrated in (13): the speaker is launching a story about events experienced by him and some company while travelling in Ireland. The story recipient is obviously not included in the reference of *we*. The inclusive use of *we* is exemplified in (14), where two siblings are talking about the origin of their dog's name, Goldie. Clancy (2010: 133) argues that, by using *we*, "[t]he siblings create an in-group, 'we the family', in opposition to the person who originally named the dog."

(13) S1: A delicacy. Mind you, I ate conga. I, I erm, I was off the west co er on a holiday to the west coast of Ireland and **we** erm **we**'re driving down the west coast and there was like some rocks and everything about, so **we** stopped and . . .

(NC: KBD-N2)

(14) SON 1: But Goldie's a girl's name like.
DAUGHTER 1: Yeah b= we didn't give her the name.

(SettCorp)

In light of this fluid nature of the origo, some researchers reject the notion of the speaker-egocentric origo and have instead proposed the notion of a sociocentric origo (e.g., Hanks 1992, 2009; Jones 1995), which "is based on the common ground shared between conversational partners" (Clancy 2010: 119).

12.2.5 Deictic categories

A number of deictic categories have been distinguished. The traditional categories are person, place, and time (Levinson 1983: 62) centered around the "deictic triad" (Caffi and Janney 1994: 366) of "here, now, and I" (Bühler 1990 [1934]: 117). Person deixis "operates on a basic three-part division, exemplified by the pronouns for first person ('I'), second-person ('you'), and third person ('he', 'she', and 'it')" (Yule 1996: 10). In cases where the underlying origo is the speaker's origo, person deictics reflect the speaker's perspective on the role of participants (speaker and addressee) and non-participants (third persons) in the speech situation (Levinson 1983: 69; Lyons 1977: 638). Place, or locational, deixis "concerns the specification of locations relative to anchorage points in the speech event" (Levinson 1983: 79), thus indexing the speaker's relationship vis-à-vis referents in terms of space. The most obvious locational deictics in English are "the adverbs 'here' and 'there' and the demonstratives 'this' and 'that'" (Fillmore 1997: 62), with both pairs encoding different degrees of proximity relative to the speaker, "here" and "this" typically associated with objects close to, and "there" and "that" to objects distant from, the speaker (Levinson 1983: 81; Hanks 2009: 11). Time, or temporal, deixis indexes the speaker's relation to referents along the temporal continuum. The inventory of time deictics not only includes adverbs such as *yesterday*, *today*, and *tomorrow*, but also time adverbials such as *last Tuesday*, *this afternoon*, *next year*. Most importantly, though, time deixis also includes tense: "tense is one of the main factors ensuring that nearly all sentences when uttered are deictically anchored to a context of utterance" (Levinson 1983: 77). Intriguingly, unlike person and place deictic systems, which may or may not be shared, "time deictic systems *always* seem to be shared" (Rühlemann 2007: 192) between the participants of the same speech situation. That is, references such as *now*, *last month*, and *tomorrow* are equally close or distal to the participants' respective origos or viewpoints.

A deictic dimension closely related to the concepts of time and space is discourse, or text, deixis: "the use of expressions within some utterance to refer to some portion of the discourse that contains the utterance (including the utterance itself)" (Levinson 1983: 85). Discourse deixis is thus "deixis in text" (Marmaridou 2000: 93). On this definition, a large number of expressions can count as discourse deixis, including the class of pronouns *provided* they are used to refer, not "to the same linguistic entity as a prior linguistic expression" (Levinson 1983: 86), but to the "linguistic expression (or chunk of discourse) itself" (Levinson 1983: 86). Consider (15) (from Lyons 1977: 667), where the pronoun *it* does not refer to the animal referred to by A but to the word used to denote it.

(15) A: That's a rhinoceros.
B: Spell **it** for me.

(Lyons 1977: 667)

Discourse deixis is also found in what has been referred to under various labels as "left dislocation" (e.g., Ashby 1988), "front dislocation" (Leech 2000), "themes" (Aijmer 1989), and "headers" (Carter and McCarthy 2006), a syntactic choice which is highly frequent in conversation while virtually absent from written prose (Biber et al. 1999: 957). An example is given in (16),

(16) [**Natalie**], [**she**]'s a right little squealer, she is. (BNC: KDM-N1)

where the full noun phrase (*Natalie*) preceding the clause is co-referential with the dummy pronoun (*she*) inside the clause. Not only do headers provide a convenient syntactic analysis "breaking up the utterance into more manageable units" (Aijmer 1989: 147), thus attending to conversationalists' "limited working memory" (Biber et al. 1999: 1067), but they also have a role in "establishing a topic" (Biber et al. 1999: 957; see also Carter and McCarthy 1995). That is, they help to structure discourse: the fronted item anticipates the theme of the forthcoming discourse (see Levinson 1983: 88).

Discourse deixis also includes so-called discourse markers, that is, expressions used to signal an utterance's or a sentence's relation to the surrounding text, most of them occurring in utterance-initial position, thus constructing for the listener the perspective in which to view the upcoming (portion of) discourse, viz. as continuation, contrast, response, elaboration, digression, transition, and so on. Utterance-initial *anyway*, for example, "seems to indicate that the utterance that contains it is not addressed to the immediately preceding discourse, but to one or more steps back" (Levinson 1983: 85). Research on discourse markers is perhaps the strand of research in which corpus linguistics has made its greatest contribution to the study of deixis. Work includes corpus analyses of the marker *like* examined in the light of relevance theory (Andersen 1998), *oh* and *ah* as signals of mental processes (Aijmer 1987), *cos* as a continuation marker (rather than a subordinating conjunction: Stenström 1998), to name only a few. Discourse deixis further concerns the use of reporting clauses such as *he said*, *she's like*, to indicate that what follows is to be attributed to some other speaker and hence some other speaker's origo (see Rühlemann 2007: 121). A wealth of corpus research has recently been concerned with the "new" quotatives GO and BE *like* (see, for example, Buchstaller 2011), two forms that invariably signal that the speaker is embarking on constructed dialogue (rather than some other form of discourse presentation such as the verb SAY, which can introduce both direct and indirect speech and thought).

Another deictic category is social deixis, which "involves the marking of social relationships in linguistic expressions, with direct or oblique reference to the social status or role of participants in the speech event" (Levinson 2004: 119). Among the axes along which these relations can be defined are the "speaker to addressee" axis instantiated, for example, in the "respectful pronouns" (Levinson 2004: 120) *Sie* (in German) or *vous* (in French) as opposed to the more familiar forms (*du* and *tu*, respectively), and the "speaker to formal setting" axis which brings distinct registers into play "used on formal occasions, where *eat* becomes *dine*, *home* becomes *residence*, etc." (Levinson 2004: 120).

The final deictic category is what Lyons (1977: 677) called "empathetic deixis," that is, indication of how the speaker is "personally involved with the entity, situation or place to which he is referring or is identifying himself with the attitude or viewpoint of the addressee." This field of deixis typically concerns choices between binary word pairs, where "'this' is selected rather than 'that', 'here' rather than 'there', and 'now' rather than 'then'" (Lyons 1977: 677). That is, more abstractly, empathetic deixis involves "a move from 'origo-farther' reference to 'origo-nearer' reference" (Rühlemann 2007: 192), a move also found in the switch, in storytelling, from Narrative Past (past tense) to Historic Present (present tense), to relate events in an anterior (past) situation, as in (17), where the narrator uses the Historic Present throughout to introduce constructed dialogue. It can be argued that this usage reflects the narrator's level of involvement with the speech exchange that the story is about.

(17) Like, Kerry goes to me, cos I've got a black kitten he **goes**, "what are you gonna call it?" I goes "dunno." Goes "call it Malcolm X." I goes "shut up!" He goes "call it, call it Ma call it Martin X" and then he says "call it Nigger." I think Nigger's a good name but, you know what I mean like "come here Nigger!" But it's, it's racist. If I'd a, if I were walking down the street going "(mimics kissing) Nigger!" [unclear] sorry! [laugh] (NC: KPG-N1)

12.3 Case study

12.3.1 Introduction

Surprisingly, given its theoretical importance, "deixis is one of the most empirically understudied core areas of pragmatics" (Levinson 2004: 97). Similarly, few corpus studies have focused explicitly on deixis to a significant extent.

In the case study we report in the present section, we examine the use of *this* in conversational storytelling. Previous research suggests that in storytelling *this* fulfills a crucial, and peculiar, role. For example, Halliday and Hasan (1976: 61) observe that "in narrative of a traditional kind, we often find *that* where, in conversational narrative, a speaker would tend to use *this*, conveying

a sense of immediacy and also of solidarity with the hearer, of shared interest and attention . . . It is the assumption of shared interest and attention which lies behind the use of the 'near' forms, *this* and *these*." Wald (1983) refers to this use of the demonstrative as "new-*this*," Biber et al. (1999: 274) term it "introductory *this*." The name implies the preference of introductory *this* to "introduce new information, especially in telling a story or introducing a new topic" (Biber et al. 1999: 274). According to Wald, the phenomenon is "widespread among vernaculars in both the United States and England and yet it is relatively new in origin" (Wald 1983: 94). Introductory *this* alternates with the indefinite article *a/an* (see Perlman 1969, Wald 1983, Rühlemann 2007), sharing with it the introduction of a referent that is "indefinite specific" (Wald 1983: 95).

Traditional accounts of the demonstratives have been based on a static model centering around the distinction between proximity (*this*) and distality (*that*). The case study is intended to contribute to a small body of evidence suggesting that the choice of *this* is determined by a variety of factors including not only proximity but also information status vis-à-vis the referent, the relationships between speaker and hearer, and so forth (see Strauss 2002).

Following Halliday and Hasan (1976), *this* can principally be used in three major reference patterns: exophoric, referring to an entity outside the text; endophoric, referring to an entity identifiable inside the text; and non-phoric, referring to an entity which "is present neither in the text nor in the situation but only in the speaker's mind" (Halliday and Hasan 1976: 61). Altogether seven subtypes can be distinguished. They are presented and illustrated in (18)–(24).

Exophoric

(18) Time deictic: Mum had a perm **this morning**

(NC: KBS-N1)

(19) Place deictic: Teachers are very unfair in **this schoo**l in n it

(NC: KNV-N2)

(20) Person deictic: What about **this man here** with a great big bass drum?

(BNC-C: KD0 11300)

Endophoric

(21) Anaphoric: the picture that comes on is Newlands Park you know. At the moment **this** is a dreadful area

(NC: KBY-N1)

(22) Cataphoric: I must just tell you **this**, Laura did make me laugh

(NC: KBG-N2)

Non-phoric

(23) Recognitional: you know **this** young girl that was killed along Benji Avenue with her mother?

(NC: KCP-N1)

(24) Introductory *this:* I had **this** octopus once in Germany and it, we 'd gone out for a meal and I was gonna have steak and mushrooms and

(NC: KBD-N1)

Given that the reader is already familiar with temporal, locational, and personal deixis and anaphoric and cataphoric reference, we will only characterize the non-phoric uses of *this* in more detail.

Both uses have in common that the reference through *this* "is present neither in the text nor in the situation but only in the speaker's mind" (Halliday and Hasan 1976: 61). In this respect, *this* in (23) and (24) can be seen as "non-phoric" (Halliday and Hasan 1976: 61; Strauss 2002). The fact that *this* attaches to a referent which is present "only in the speaker's mind" entails that reference through recognitional and introductory *this* is "first-mention," or discourse-new, reference. The differences are twofold. First, the reference of *this young girl* in (23) is "recognitional" in the sense that it is "used to activate shared knowledge" (Diessel 1999: 105): the speaker uses *this* to call into the consciousness of the hearer a referent both participants have access to. (Note the use of question intonation as indicated by the question mark in the example, indicating the speaker's uncertainty whether the interlocutor will recognize the intended referent.) That is, recognitional *this* facilitates hearer-old reference. By contrast, the reference of *this octopus*, which occurs at the launch of a story whose events happened decades ago, is obviously not shared between narrator and listener; the reference is hearer-new. Second, because in (23) the reference is identifiable for the hearer via shared knowledge, the referent of *this young girl* is unlikely to be mentioned again in the subsequent discourse. Conversely, the reference of *this octopus* is "introductory" in the sense that its referent is what the whole story is about and therefore likely to receive subsequent mentions as the story unfolds. So, while both examples illustrate non-phoric (discourse-new) reference, (23) is an instance of recognitional (hearer-old) reference, while (24) is an illustration of introductory (hearer-new) reference.

The two functions have received only scant attention in the literature (e.g., Halliday and Hasan 1976, Diessel 1999, Biber et al. 1999, Rühlemann 2007). One of the very few quantitative studies is Strauss (2002). Building on Halliday and Hasan (1976), Strauss distinguishes between basic reference patterns with *this*: anaphoric, cataphoric, exophoric, and non-phoric and examines how

these functions are distributed across four registers (lecture, radio discussion, television interview, and conversation). Strauss's study invites follow-up analysis on three counts. The study is based on a very small 45,000-word corpus. Second, the two non-phoric subfunctions of *this* are not distinguished and, as a consequence, not investigated in terms of their association with genre. Finally, the intricate association of introductory *this* with conversational narrative is not considered in any depth.

If introductory *this* characteristically serves to draw attention to information or a topic which is *new* to the discourse, it would not be surprising if it were also "introductory" in a *positional* sense: in the sense of demonstrating a preference for occurring in early sections of the story. That is, a key feature of introductory *this* might be its position within the storytelling process. Recent work has demonstrated the close association of lexis with positions in texts (e.g., Hoey 2005, Scott and Tribble 2006, Mahlberg and O'Donnell 2008, Hoey and O'Donnell 2008, Römer 2010, Rühlemann et al. 2011, O'Donnell et al. 2012, Hoey and O'Donnell 2013, Rühlemann 2013). In Hoey's theory of "lexical priming," the claim is made that "every word is primed to occur in, or avoid, certain positions within the discourse" (Hoey 2005: 13). This claim is referred to as "textual colligation." Also, textual positioning is "likely to be genre- and even subgenre-specific" (Hoey and O'Donnell 2008: 300). Most analyses so far have concentrated on textual colligation in written genres; the first detailed examination of positioning in speech is Rühlemann's (2013) study of interjections as quotation markers in narrative.

So, the specific aims in this case study are twofold: first, we wish to establish whether *this*, used as introductory *this*, is "primed" to occur in early positions within storytelling; second, we wish to examine introductory *this* in terms of the above-noted reference patterns.

12.3.2 Data and methods

The data used come from the Narrative Corpus, henceforth NC, a corpus of conversational narratives and their surrounding non-narrative contexts. The texts collected in the NC are extracted from the conversational component of the British National Corpus (BNC-C) and retain all the XML annotation used in that "mother corpus." This pre-existing annotation is richly augmented by several layers of annotation intended to capture narrative-specific characteristics. The additional annotations capture, for example, participant status (different types of narrator and recipient), quotatives (*says*, *went*, BE *like*, etc.), and discourse presentation (direct, indirect, etc.), to name only a few. See Rühlemann and O'Donnell 2012 for more detail on the NC's annotation. One further type of annotation, which is central to the present analysis, is the markup of

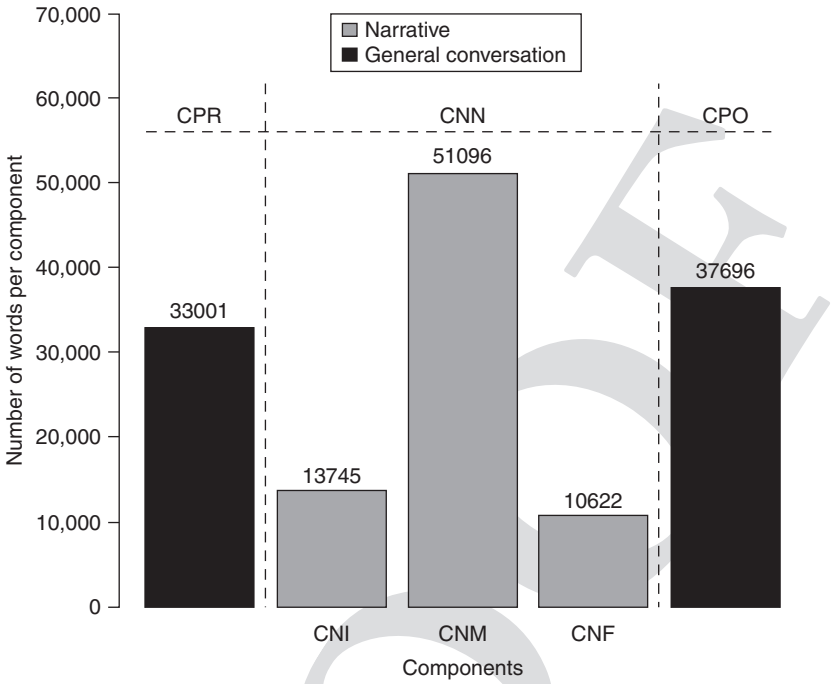

Figure 12.1 Componential structure of texts in the Narrative Corpus.

textual components, shown in Figure 12.1 (the numbers above the bars indicate the number of words in the components).

Texts in the NC contain macro- as well as micro-components. The macro-components comprise general conversation preceding the story (pre-narrative component, CPR), the storytelling itself (narrative component, CNN), and general conversation following the story (post-narrative component, CPO), while the micro-components include the story-initial utterance (narrative-initial component, CNI) and the utterance which concludes the storytelling (narrative-final component, CNF); utterances between CNI and CNF are referred to (but not annotated as such) as the narrative-medial component (CNM). Thanks to the componential annotation, these components can be addressed and analyzed as subcorpora.

A lexical item's textual position is calculated as the proportion of number of words preceding the item out of the total number of words in the text. Based on this procedure, positional values range between 0 and < 1. For example, to return to excerpt (7), *this* in *this octopus* occurs as the third word in a story (CNN) with altogether 108 words. Thus, the position of *this* in the text

is 2/108 = 0.02. To retrieve positions, XQuery (see Walmsley 2007) was used, a query language for XML texts based on XPath (see Clark and DeRose 1999). XQuery allows the researcher to navigate the hierarchical tree structure of XML texts, select elements of interest and extract the results (for a more detailed description of XPath and XQuery, see Rühlemann 2013: Chapter 2).

Moreover, we made use of keyness analysis (see Scott 1997, Scott and Tribble 2006) to identify items with preferences (or "primings") for occurrence in textual positions as defined by the component annotation. Items such as words, lemmas, *n*-grams, POS-tags and semantic domains (see Rayson 2008) are considered key if they occur in a given text/text type with an "unusual frequency in comparison with a reference corpus of a suitable kind" (Scott and Tribble 2006: 55). Keyness analyses can give an indication both of the "aboutness" of a text/text type and of its stylistic characteristics (see Scott and Tribble 2006: Chapter 4). O'Donnell et al. (2012) have introduced intra-textual keyness analysis, where keyness is calculated for lexical items in a subcorpus/subset of a text/text type as compared to another subcorpus/subset of the same text/text type. We adopt this method here to identify items associated with specific narrative components.

12.3.3 Results

A first approximation to the distribution of *this* across textual components in the NC is shown in Figure 12.2, which depicts density lines for *this* in the pre-narrative (CPR), the narrative (CNN), and the post-narrative (CPO) components. The curves represent the estimated probability densities of occurrence of *this* and *these* in the three textual components CPR, CNN, and CPO. Density curves can be thought of as "smoothed histograms" (Baayen 2008: 25) with density being "a measure of the relative probability of 'getting a value close to *x*'" (Dalgaard 2008: 59). For the singular form *this*, it can be seen that the probabilities are highest in the early positional intervals in narrative (CNN, grey line), more specifically in the second positional interval (between 0.1 and 0.2), where the probabilities for *this* form a big hump. Apart from that early peak, the estimated densities curves for all three components follow remarkably similar trajectories. For the plural form *these*, by contrast, no such early maximum can be observed in CNN; rather, a maximum is found around the value 0.6 for pre-narrative (CPR) components. While highly suggestive and helpful for highlighting the presence of introductory *this* in CNN, it must be noted that for both *this* and *these*, the differences between the distributions across the three components are not enough to warrant significance. According to two-sided Kolmogorov–Smirnov tests (Baayen 2008: 79; Gries 2009: 163–165), the *p*-values are below the threshold level of 0.05. That is, we cannot assume the distributions are different. If we suspect that *this* fulfills a

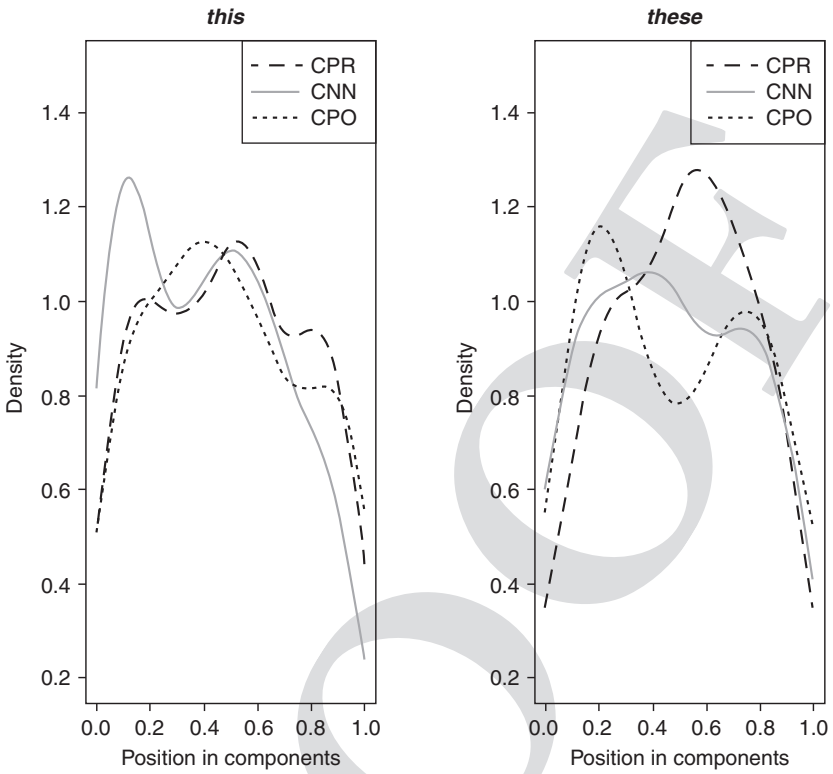

Figure 12.2 Comparison of estimated probability densities for positions of *this* and *these* in textual components (CPR, pre-narrative; CNN, narrative; CPO, post-narrative).

special role in narrative and that it does so early in the storytelling process, investigating its positional distribution across the narrative *as a whole* is too crude a method.

The NC's componential structure provides a convenient starting point to refine the method. It allows us to count the number of occurrences of *this* and *these* across not only the macro- but also the micro-components: narrative-initial (CNI), narrative-medial (CNM), and narrative-final (CNF) utterances.

As shown by Table 12.1, *this* clearly prefers narrative-initial utterances (CNI). While the frequencies of *this* are stable across the two non-narrative components CPR (pre-narrative) and CPO (post-narrative) with around 4 occurrences per 1,000 words, the number for narrative-initial utterances (CNI) is twice as high (8 occurrences). From this high, the numbers drop

Table 12.1 *Distribution of* this *and* these *across textual components (RF, raw frequency; NF, normalized frequency, per 1,000 words).*

	this		these	
Component	RF	NF	RF	NF
CPR	138	4.18	27	0.82
CNI	110	8.00	20	1.46
CNM	252	4.93	58	1.14
CNF	36	3.39	7	0.66
CPO	175	4.64	35	0.93

steadily across narrative-medial utterances (CNM: 4.93) to narrative-final utterances (CNF: 3.39), where the frequency is even lower than in the non-narrative stretches. For *these*, too, we observe the highest normalized frequency in CNI (1.46), but the leap from CPR to CNI and the drop from CNI to the following components is much less dramatic than for *this*. Are these differences due to chance or a reflection of the "true" differences in the population (conversational narrative)? According to chi-squared tests, the differences for *this* are very highly significant ($\chi^2 = 36.49$, d.f. $= 4$, $p = 2.295 \times 10^{-7}$), whereas the differences for *these* are insignificant ($\chi^2 = 6.29$, d.f. $= 4$, $p = 0.18$). That is to say that there is a significant association between *this*, but not *these*, and textual component. Given the insignificant result for *these*, we will exclude the plural form from the remainder of the analysis. While a chi-squared test "does not allow one to make cause and effect claims" (Oakes 1998: 24), we would like to know how the variables contribute to the significant result. To answer this question, the Pearson residuals are a useful diagnostic. They "indicate the degree to which the observed frequencies deviate from the expected ones" (Gries 2009: 197).

Table 12.2 reveals that the significant result is mostly "powered" by the observed frequencies in CNI, which greatly exceed the expected frequencies and where the deviation is hence greatest (5.28), and in CNF, where far greater frequencies were expected than observed (–2.20).

These results suggest two findings: first, *this* is primed to occur in narrative-initial position; second, it is primed to avoid occurring in narrative-final position. That *this* is negatively primed not to occur in story-closing utterances is also suggested by the fact that *this* is among the most strongly negative key words in narrative-final components (CNF) compared to narrative components (CNN) minus CNF (LL $= 8.633$, $p < 0.01$).

Since *this* is not any lexical item but a core referring expression, an analysis of the reference patterns it enters into will get us closer to understanding why

Table 12.2 *Pearson residuals for the distribution of* this *across textual components.*

	CPR	CNI	CNM	CNF	CPO
this	–1.78	5.28	0.23	–2.20	–0.62
All other words	0.12	–0.37	–0.02	0.15	0.04

the positive association of *this* with story beginnings and negative association with story endings may exist.

To approach the question of how *this* is actually used in narrative-initial utterances, we queried the NC for key words and key 2-grams and 3-grams in narrative-initial utterances (CNI) as compared to all other utterances in the stories (CNM and CNF): see Table 12.3. Not surprisingly, considering the above-noted significant over-representation of *this* in narrative-initial utterances (CNI), we find the demonstrative among the top key words in CNI. More surprisingly, *this* also figures among the key 2-grams (*this morning*, ranked third) and key 3-grams (*there's this*, ranked 16th). The association with the noun *morning* is of particular relevance in the present connection in that the phrase *this morning* is clearly a time deictic expression meaning the first section of "the diurnal span in which the speaking event takes place" (Levinson 2004: 114). This use of *this* is, then, obviously not the introductory *this* we are looking for and for which we claim a special role. A better candidate for introductory *this* is the key 3-gram *there's this*, which may more readily attract complementation by nouns indicating new information or new topics. To examine this possibility, and the functions of *this* in narrative-initial utterances in general, a qualitative concordance analysis was carried out; that is, all 110 instances of *this* in CNI were inspected in the context of their utterances and, where necessary, of the story.

This analysis involved four steps: (i) code occurrences of *this* in CNI for reference type, (ii) calculate nouns in R1–4 position (that is, within the span of one to four words on the right of *this*), (iii) assign the collocating nouns to semantic groups, and (iv) double-code each *this* in CNI for both reference type and semantic group of collocate.

Step (i) (coding for reference type) produced the following results, shown in Figure 12.3. Instances of *this* were categorized as U (unclear) in restarts occurring right after *this*, as in *Er, just* **this *this*** *woman came round to the door last night* (KD7-N1), where the first *this* was labeled U while the second was labeled I (introductory). Another important component of the uses set to U were occurrences within direct speech, as in *Yeah, well he was, when he was*

Table 12.3 *Key items in narrative-initial utterances (CNI) (Freq. A) compared to narrative-medial (CNM) and narrative-final utterances (CNF) (Freq. B).*

	Key words				Key 2-grams				Key 3-grams			
N	Item	Freq. A	Freq. B	LL	Item	Freq. A	Freq. B	LL	Item	Freq. A	Freq. B	LL
1	*last*	53	61	47.57	*used to*	30	33	28.25	*used to go*	8	1	21.37
2	*today*	19	14	25.36	*last night*	21	16	27.34	*you know what*	13	7	21.19
3	*morning*	28	34	23.67	*this morning*	16	8	27.16	*we went to*	11	5	19.60
4	*used*	30	39	23.38	*was on*	15	7	26.38	*when I was*	9	5	14.41
5	*this*	110	280	23.24	*my dad*	9	4	16.21	*she's got*	13	13	13.46
6	*yesterday*	21	21	21.75	*know what*	22	33	14.17	*the other day*	7	3	12.83
7	*once*	15	10	21.46	*I went*	26	45	13.37	*do you remember*	6	2	12.24
8	*remember*	32	49	20.00	*other day*	7	3	12.83	*I tell you*	6	2	12.24
9	*night*	38	67	18.92	*had a*	27	50	12.30	*it was on*	6	2	12.24
10	*when*	94	243	18.91	*mean we*	5	1	12.06	*did you see*	5	1	12.03
11	*one*	90	233	18.02	*last week*	11	11	11.40	*I mean we*	5	1	12.03
12	*other*	32	55	16.65	*the other*	22	38	11.35	*I used to*	7	4	11.03
13	*went*	76	192	16.44	*at school*	7	4	11.03	*do you know*	12	14	10.61
14	*holiday*	6	1	15.10	*I used*	7	4	11.03	*what I mean*	7	5	9.55
15	*we*	133	406	13.93	*you remember*	7	4	11.03	*the other night*	5	2	9.46
16	*goes*	46	104	13.57	*'s like*	15	21	10.63	*there's this*	5	2	9.46
17	*school*	18	25	12.89	*I had*	28	57	10.55	*and he goes*	8	8	8.28
18	*about*	77	211	12.67	*I tell*	6	3	10.18	*know what I*	8	8	8.28
19	*Ryan*	10	8	12.55	*other night*	5	2	9.46	*they used to*	5	3	7.65
20	*up*	100	299	11.54	*was about*	5	2	9.46	*the last time*	6	5	7.29

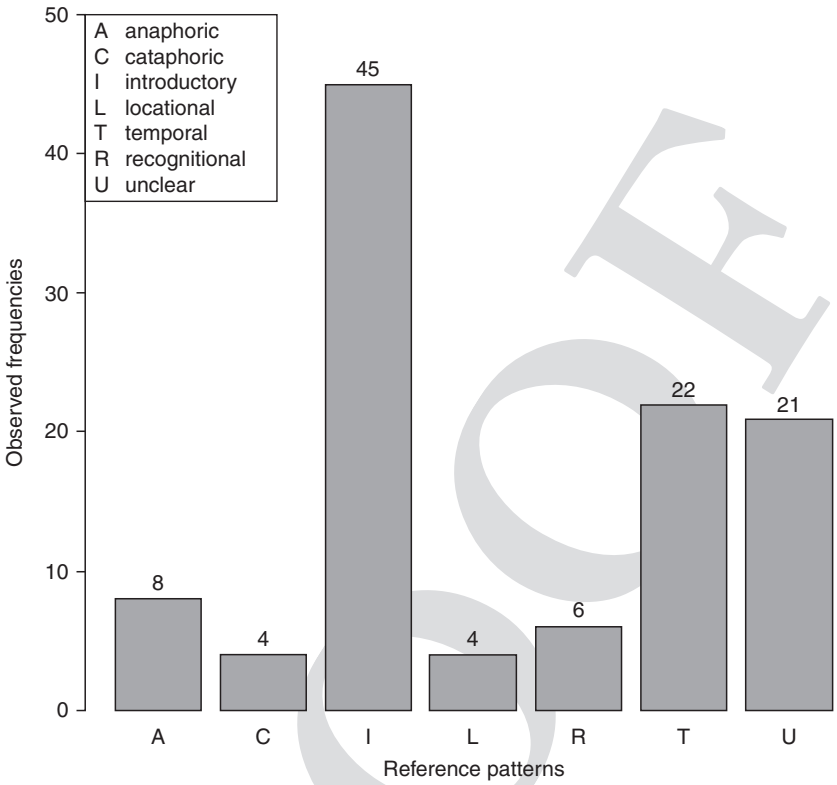

Figure 12.3 Frequencies of reference patterns of *this* in narrative-initial utterances (CNI).

littler his dad was sort of saying^{QSG} [^{MDD}*one day you'll have all* **this**] (KDJ-N2), where the reference is the reported speaker's reference, not the narrator's.

This in story-launching utterances predominantly performs the function of introductory *this* (45 occurrences, 41%). The second most commonly used reference pattern is as a component of time deictic expressions (22 occurrences, 20%). The U (unclear) category is large too (21 occurrences, 19%), mainly due to occurrences of *this* within direct speech (12 occurrences). While all other reference patterns see one-digit frequencies, no occurrence of *this* as a personal deictic referring to a person in the situational (that is, storytelling) context was found in the data.

In determining the nouns that *this* in CNI collocates with (step (ii)) and assigning them to semantic groups (step (iii)), we found four groups, as shown in Table 12.4. The overwhelming majority of instances of *this* in narrative-initial

Table 12.4 *Semantic groups of nouns collocating with* this *in R1–4 in narrative-initial utterances (CNI).*

Semantic group	People	Time	Object	Location	Non-adnominal	Unclear	Total
Total frequency in R1–4	26	22	22	11	25	4	110

utterances (CNI) were used adnominally, that is, as a demonstrative pre-modifying a noun, as in ***this bloke*** *he cracked up* (NC: KC2-N1); only 25 instances, or 23%, were non-adnominal, as in ***this****'ll be his fourth week* (NC: KCT-N2). The largest adnominal group is the "people" group (26 occurrences). This group comprises nouns denoting human referents such as *girl*, *man*, *bloke*, *woman*, and *lad*, to name the five most frequent forms (in descending order). There are two second-largest groups (22 occurrences each), one including nouns encoding "time" meanings such as *morning*, *time*, *week*, *year*, and *afternoon*, and the other comprising "objects," such as *piece*, *rubbish*, *hat*, and *stick*. Finally, the least frequent group was made up of nouns denoting "location" (13 occurrences), such as *mountain*, *school*, *area*, and *Germany*.

Finally, in step (iv), all instances of *this* in CNI were coded both for reference pattern and semantic group. The breakdown is shown in Figure 12.4. A number of observations can be made. The most frequent category is I_p (Introductory_people), where *this* is used as introductory *this* and combines with a "people" word, thus constructing discourse-introductory reference to a story character (25 occurrences). The second most frequent category is T_t (Temporal_time), where *this* is used as a time deictic and collocates with a noun denoting "time" (22 occurrences). This category is also noteworthy in that temporal reference with *this* is made solely in collocation with "time" nouns. The third most frequent category is once again an introductory *this* category, with 13 occurrences of *this* in combination with an "object" noun, thus referring to a non-human referent in the story. All other categories but U_n (Unclear_non-adnominal) have single-digit frequencies.

12.3.4 Discussion

The results reported so far are relatively straightforward. As regards *this*, *tout court* (regardless of what reference pattern it enters into), we found that the distribution of *this* (and also *these*) was skewed towards early positions in storytelling. Specifically, we noted a powerful attraction of the singular form *this* (but not of the plural form *these*) to the first utterance in a story (narrative-initial utterance, CNI): there, *this* occurs with significantly higher frequency than

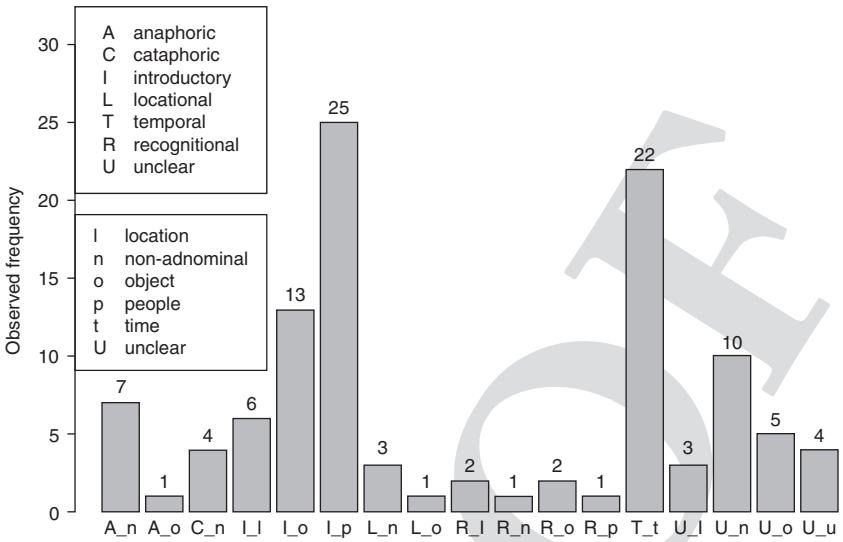

Figure 12.4 Breakdown of frequencies of reference patterns for *this* in CNI by semantic group.

in utterances in the middle of the story (CNM), let alone story-final utterances (CNF), where it is found least. This unequal distribution across utterances in storytelling suggests, in Hoey's (2005) priming terms, that *this* is primed to occur in story-initial utterances and primed to avoid occurring in story-final utterances.

Further, as regards reference, we found that in narrative-initial utterances adnominal *this* is predominantly used as introductory *this*, introducing a referent not yet mentioned in the storytelling discourse. The second most frequent reference type through adnominal *this* was as a time deictic, indexing the told time (the time the story events happened) to the telling time (what Levinson 1983 refers to as "coding time"). Finally, we saw that introductory *this* most frequently introduces human referents ("people"), but also, to some extent, non-human referents ("object"); this latter observation supports Wald's (1983: 102) finding that use of "new-*this* extends to inanimates." These two most important uses of introductory *this* are illustrated in the following excerpts. In (25)–(27), we present three examples of introductory *this* introducing human referents, and, in (28)–(29), two examples introducing "object" referents. In all five excerpts, we highlight not only the occurrence of introductory *this* (using arrows) but also the subsequent mentions of the same referent introduced by *this* (using bold face). Inspection of the contexts preceding the uses of introductory *this* in the excerpts confirms that the references are invariably "first-mentions."

In (25), the girl that the narrator is referring to as *this girl* in CNI is mentioned again no less than nine times in the story (see the anaphoric pronominal references to her). Evidently, the story revolves around the referent as the story protagonist.

(25) "Waitress" (Type: T10 / Embed Level: EC2)

	CNI		
→	S3	PNP	Eh! Our Arthur, Arthur sat there and **this girl** come to clear the pots away and **she**'s been round lots of tables, you know, collecting the cups up together and **she** comes in and **she** goes ^{QGZ} [^{MDD} ooh!] **She** says ^{QSZ}, [^{MDD} it smells.]
	CNM		
	S2	PRC	(laughing) That's very nice!
	S3	PNP	And I looked at **her** and I thought ^{QTD} [^{MDD} you can't say that!] And Alan looked at **her** and then **she** looked, **she** goes ^{QGZ}, [^{MDD} oh I meant,] **she** says ^{QSZ}, [^{MDD} you smile nice.] (laughing) He said ^{QSZ}, [^{MDD} don't get close].
	S2	PRR	Yeah.
	CNF		
	S3	PNP	[^{MDF} Ooh! You smell.]
			(NC: KCR-N1)

That the referent marked by introductory *this* is central to the unfolding story transpires also from (26), where the referent of *this quite chubby bloke* is the sole explicit protagonist in the story.

(26) "Chubby bloke" (Type: T10 / Embed Level: EC3)

	CNN		
	CNI		
	S2	PNP	There was one at our yeah. Exactly the same as ours. When we was about fourteen in the fourth year at school, there
→			was **this quite chubby bloke** [unclear] **he**'d have a towel just there while you shower.
	CNM		
	S2	PRC	Yeah that's right, that's what [unclear]
	S3	PXX	[unclear]
	S2	PNP	**He**'d sort of hang out, **he**'d [unclear] shower with **his** [unclear]
	S2	PRR	(laugh)
	S2	PRC	you know **he**'d go ^{QGB} like that, [^{MDD}yeah look at me I'm, I'm mature] sort of thing.
	S2	PRC	That's so funny.
	CNF		
	S2	PNP	It wasn't at the time because you had a real complex about it but

Although, in (27), the reference to *this old farmer* is only repeated once in the surface structure, the referent is still the leading figure in the story – what Wald (1983: 104) refers to as "implicit recurrence" – in that without the farmer's advice the teller and his/her company would not have come to the realization that fishing in the canal is useless.

(27) "Drained canal" (Type: T10 / Embed Level: EC2)

	CNN		
	CNI		
	S1	PNS	I don't know what's going on. I don't we got about three, three thirty in the morning, both of them went out to er canal somewhere up Dulgate, past Dulgate we set up and we'd we'd been fishing for about two and half hours it's aba about
→			six thirty in the morning **this old farmer** comes up says ^{QSZ} [^{MDD}er aye, aye lads,] **he** said ^{QSD} [^{MDD}er I wouldn't bother it, they drained this area of the canal a few (laughing) months ago!] And we said^{QSD} [^{MDD}, oh!]
	S2	PRR	Yeah. (laugh)
	CNF		
	S1	PNS	(laugh) (laughing) Sat there watching our floats for hours! I mean
	S1	PRR	Yeah.

As regards inanimate object references through introductory *this*, a similarly central role for the referent can be observed. In (28), the above-discussed octopus soup figures in the narrator's story as the central entity around which the actions by the human referents revolve. Here, too, the semantic centrality of the object shines through in the surface structure where the reference is repeated a number of times.

(28) "Octopus soup" (Type: T10 / Embed Level: ES)

	CNN		
	CNI		
→	S2	PNP	I had **this octopus** once in Germany and it, we'd gone out for a meal and I was gonna have steak and mushrooms and
	S1	PRR	Mm.
	CNM		
	S2	PNP	er, the lads said ^{QSD} [^{MDD}what are we having the starters like?] I said^{QSD}, [^{MDD}oh I'll have what you're ordering.] Well they ordered soup of the day you see and they brought **this**, I tasted **it** had a real funny, a weird taste
	S1	PRR	Mm.
	S2	PNP	anyway, and I sort of said ^{QSD} to the waitress, [^{MDD}why what was it?] She said ^{QSD} [^{MDD}octopus.] [^{MDF}Urgh!]

S1	PRC	(laugh) You wouldn't have had **it** if you'd have known?
CNF		
S2	PNP	No. I don't think I would.
S1	PRR	No.

(NC: KBD-N1)

In (29), by contrast, *this almighty swing* introduced in CNI is not referred to again in the subsequent storytelling and, it seems, it does not occupy the center of the semantic substructure of the story. Rather, that center is taken over by the golf club which, due to the *almighty swing*'s thrust, lands in a tree. Excerpt (29), then, illustrates that introductory *this* may attach not only to story-central but also peri-central entities.

(29) "Flying six iron" (Type: T10 / Embed Level: ES)

	CNN		
	CNI		
	S1	PNS	Had to laugh yesterday towards the end of er our round, it was really tipping down, misty, fog,
→			couldn't see a hundred yards I had **this almighty s** I said **QSD** [MDD well I'll let the six iron,] I did **this almighty swing** the ball went one way and the club went right up the top of a tree. It must've been fifteen foot at least up in the air teetering on a top branch.
	CNM		
	S2	PRR	(laugh)
	S1	PNS	I thought QTD [MDD oh my god, lost me six iron,] all of a sudden [MDF crash], it came down.
	S2	PXX	(???)
	CNF		
	S1	PNS	Old Louis he didn't know which way anything had gone. He said QSD [MDD well I couldn't see any of that didn't even see the club go up the tree.]

(NC: KC1-N2)

The core function, then, of introductory *this* is to introduce entities that are going to play a pivotal role in the unfolding narrative; for similar observations see Wald (1983). Introductory *this* marks (peri-)central, or proto-, referents. This is an important observation, on two counts. First, it allows us to refine the above-noted textual colligation claim: narrators are not only primed to use *this* in narrative-initial utterances, but, more specifically, they are primed to use it as introductory *this*, that is, to introduce key referents. So we can add an additional textual priming statement: when it occurs in

narrative initial position, *this* is primed to function as a marker of the key narrative referent. This kind of priming is referred as a textual semantic association (Hoey 2005, Hoey and O'Donnell 2013). Second, given its intimate association with (peri-)central entities, introductory *this* acts as a type of "theme marker," alerting the addressee to the entities' centrality, or, as Jucker and Smith (2003: 409) put it, "instruct[ing] the addressee to pay attention to these entities" (see also Wald 1983: 97, Smith et al. 2005: 1886). Introductory *this* can thus be seen as a form of discourse deixis (see Levinson 1983: 89). As a discourse-deictic theme marker, introductory *this* fulfills an important function in the narrator's "recipient design"[2] (Sacks et al. 1974): in highlighting who (or what) is going to be the story's key referent, it addresses the recipient's processibility needs, aiding them in constructing a mental image of the story world (see Biber et al. 1999: 265; see also Halliday and Hasan 1976: 61). The theme marking role of introductory *this* also provides a useful explanation for the textual colligations observed above, viz. the positive attraction to story-early positions, in particular story-initial utterances, and the drop in frequency in narrative-medial utterances and particularly in narrative-final utterances, where speakers seem primed to avoid using *this*: if the core function of *this* is to introduce story-central entities, story beginnings, where the scene is being set for the events to be narrated, are a much more adequate location than story endings, where the scene is being closed.

12.4 Conclusions

The case study provides strong support for the notion that textual colligation is at work not only in writing but also in spoken communication. Building on Rühlemann (2013), the first investigation into textual positioning in conversational narrative, our case study is the second to discover priming in storytelling: speakers are primed to use *this*, in its functional incarnation as introductory *this*, in narrative-initial utterances and avoid using the demonstrative in narrative-final utterances.

The case study also suggests the possibility that textual colligation in the use of *this* is genre-specific (see Hoey 2005, Hoey and O'Donnell 2013). Although our investigation has focused on conversational narrative, it appears unlikely that *this* should exhibit the same textual colligation in other spoken genres. We

[2] By "recipient design" Sacks et al. "refer to a multitude of respects in which the talk by a party in a conversation is constructed or designed in ways which display an orientation and sensitivity to the particular other(s) who are the co-participants" (Sacks et al. 1974: 272). See also the discussion and analyses in Rühlemann (2013).

base this assumption specifically on functional grounds: the clear preference for speakers to use *this* as introductory *this* is specifically adapted to needs arising in storytelling, in particular the listener's need for enhanced processibility by means of highlighting the proto-referent around which to construct the story world. Given this adaptedness to storytelling, we assume that introductory *this* will be far from typical in other spoken genres. Considering the non-comparative nature of the case study, though, this assumption is, at present, merely a hypothesis which needs to be verified in future research. If it could be verified, that is, if a core deictic such as *this* could be shown to vary significantly both in frequency and function across genres, the widely held assumption of a general "meaning" of deictic items which remains stable across genres would have to be decisively re-evaluated (see Jones 1995: 48).

Finally, the case study highlights the usefulness of developing and working with specialized corpora, such as the NC, which offer multiple layers of sophisticated annotation. The analyses performed and the findings presented here would have been impossible in traditional corpora with no annotation or POS annotation only, because even POS annotation is still "limited annotation" (Leech 2007: 134), greatly facilitating the study of lexical patterning but seriously limiting the study of discourse and pragmatic phenomena. If corpus linguistics is to contribute to a deeper understanding of context-intensive phenomena such as deixis, or most other pragmatic phenomena, the help of additional layers of annotation will be indispensable, particularly if that deeper understanding is to be gained not only from qualitative but also from quantitative analyses.

REFERENCES

Aijmer, K. 1987. *Oh* and *Ah* in English conversation. In W. Meijs (ed.), *Corpus Linguistics and Beyond*, pp. 61–68. Amsterdam and Atlanta: Rodopi.

1989. Themes and tails: The discourse functions of dislocated elements. *Nordic Journal of Linguistics* 12: 137–154.

Andersen, G. 1998. The pragmatic marker *like* from a relevance-theoretic perspective. In A. H. Jucker and Y. Ziv (eds), *Discourse Markers: Descriptions and Theory*, pp. 147–170. Amsterdam and Philadelphia: John Benjamins.

Ashby, W. J. 1988. The syntax, pragmatics, and sociolinguistics of left- and right-dislocation in French. *Lingua* 75: 203–229.

Baayen, H. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.

Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, E. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education.

Buchstaller, I. 2011. Quotations across the generations: A multivariate analysis of speech and thought introducers across 5 decades of Tyneside speech. *Corpus Linguistics and Linguistic Theory* 7 (1): 59–92

Bühler, K. 1990 [1934]. *Theory of Language: The Representational Function of Language*. Amsterdam and Philadelphia: John Benjamins.
Caffi, C. and Janney, R. W. 1994. Towards a pragmatics of emotive communication. *Journal of Pragmatics* 22: 325–373.
Carter, R. A. and McCarthy, M. J. 1995. Grammar and the spoken language. *Applied Linguistics* 16 (2): 141–158.
2006. *Cambridge Grammar of English*. Cambridge: Cambridge University Press.
Clancy, B. 2010. 'Hurry up baby son all the boys is finished with their breakfast': A sociopragmatic analysis of Irish settled and traveller family discourse. Unpublished doctoral thesis, Mary Immaculate College, University of Limerick, Ireland.
Clark, J. and DeRose, S. 1999. XML *Path Language XPath Version 1.0*. Online at: www.w3.org/TR/xpath (accessed May 2011).
Crystal, D. 2003. *A Dictionary of Linguistics and Phonetics* (fifth edition). Oxford: Blackwell.
Dalgaard, P. 2008. *Introductory Statistics with R*, second edition. Berlin: Springer.
Diessel, H. 1999. *Demonstratives: Form, Function, and Grammaticalization*. Amsterdam and Philadelphia: John Benjamins.
Fillmore, C. J. 1997. *Lectures on Deixis*. Stanford, CA: CSLI Publications.
Goffman, E. 1974. *Frame Analysis: An Essay on the Social Organization of Experience*. New York: Harper and Row.
Goodwin, C. and Duranti, A. 1992. Rethinking context: An introduction. In A. Duranti and C. Goodwin (eds), *Rethinking Context: Language as an Interactive Phenomenon*, pp. 1–42. Cambridge: Cambridge University Press.
Gries, S. T. 2009. *Statistics for Linguistics with R: A Practical Introduction*. Berlin: Mouton de Gruyter.
Halliday, M. A. K. and Hasan, R. 1976. *Cohesion in English*. London: Longman.
Hanks, W. F. 1992. The indexical ground of deictic reference. In A. Duranti and C. Goodwin (eds), *Rethinking Context: Language as an Interactive Phenomenon*, pp. 43–76. Cambridge: Cambridge University Press.
2009. Fieldwork on deixis. *Journal of Pragmatics* 41: 10–24.
2011. Deixis and indexicality. In W. Bublitz and N. R. Norrick (eds), *Foundations of Pragmatics*, vol. 1 of Handbooks of Pragmatics, pp. 315–346. Berlin: Mouton de Gruyter.
Hoey, M. 2005. *Lexical Priming: A New Theory of Words and Language*. London and New York: Routledge.
Hoey, M. and O'Donnell, M. B. 2008. Lexicography, grammar, and textual position. *International Journal of Lexicography* 21 (3): 293–309.
2013. Examining associations between lexis and textual position in hard news stories. In M. Charles, N. Groom and S. John (eds), *Corpora, Grammar, Text and Discourse: In Honour of Susan Hunston*. Amsterdam and Philadelphia: John Benjamins.
Hundt, M., Nesselhauf, N. and Biewer, C. (eds) 2007. *Corpus Linguistics and the Web*. Amsterdam and New York: Rodopi.
Jones, P. 1995. Philosophical and theoretical issues in the study of deixis: A critique of the standard account. In K. Green (ed.), *New Essays in Deixis: Discourse, Narrative, Literature*, pp. 27–48. Amsterdam and Atlanta: Rodopi.
Jucker, A. H. and Smith, S. W. 2003. Reference assignment as a communicative task: Differences between native speakers, ESL- and EFL-speakers. In E. Mengel,

H.-J. Schmid and M. Steppat (eds), *Proceedings of Anglistentag 2002, Bayreuth*, pp. 401–410. Trier: Wissenschaftlicher Verlag.

Leech, G. 2000. Grammars of spoken English: New outcomes of corpus oriented research. *Language Learning* 50 (4): 675–724.

2007. New resources, or just better old ones? The holy grail of representativeness. In M. Hundt, N. Nesselhauf and C. Biewer (eds), *Corpus Linguistics and the Web*, pp. 133–150. Amsterdam and New York: Rodopi.

Levinson, S. C. 1983. *Pragmatics*. Cambridge: Cambridge University Press.

2004. Deixis. In L. R. Horn and G. Ward (eds), *The Handbook of Pragmatics*, pp. 97–121. Malden, MA: Blackwell.

Lyons, J. 1977. *Semantics*, vols I, II. Cambridge: Cambridge University Press.

Mahlberg, M. and O'Donnell, M. B. 2008. A fresh view of the structure of hard news stories. In *Proceedings of the 19th European Systemic Functional Linguistics Conference and Workshop*. Online at: scidok.sulb.uni-saarland.de/volltexte/2008/1700/.

Marmaridou, S. S. A. 2000. *Pragmatic Meaning and Cognition*. Amsterdam and Philadelphia: John Benjamins.

Mindt, D. 2000. *An Empirical Grammar of the English Verb System*. Berlin: Cornelsen.

Oakes, M. P. 1998. *Statistics for Corpus Linguistics*. Edinburgh: Edinburgh University Press.

O'Donnell, M. B., Scott, M., Mahlberg, M. and Hoey, M. 2012. Exploring text-initial words, clusters and concgrams in a newspaper corpus. *Corpus Linguistics and Linguistic Theory* (special issue) 8 (1): 73–101.

O'Keeffe, A., Clancy, B. and Adolphs, S. 2011. *Introducing Pragmatics in Use*. London and New York: Routledge.

Perlman, A. 1969. 'This' as a third article in American English? *American Speech* 44 (1): 76–80.

Rayson, P. 2008. From key words to key semantic domains. *International Journal of Corpus Linguistics* 13 (4): 519–549.

Roberts, C. 2004. Context in dynamic interpretation. In L. R. Horn and G. Ward (eds), *The Handbook of Pragmatics*, pp. 197–220. Malden, MA: Blackwell.

Römer, U. 2010. Establishing the phraseological profile of a text type: The construction of meaning in academic book reviews. *English Text Construction* 3 (1): 95–119.

Rühlemann. C. 2007. *Conversation in Context: A Corpus-Driven Approach*. London: Continuum.

2013. *Narrative in English Conversation: A Corpus Analysis of Storytelling*. Cambridge: Cambridge University Press.

Rühlemann, C, Bagoutdinov, A. and O'Donnell, M. B. 2011. Windows on the mind: Pauses in conversational narrative. *International Journal of Corpus Linguistics* 16 (2): 198–232.

Rühlemann, C. and O'Donnell, M. B. 2012. Towards a corpus of conversational narrative: Construction and annotation of the *Narrative Corpus*. *Corpus Linguistics and Linguistic Theory* 8 (2): 313–350.

Sacks, H. 1992. *Lectures on Conversation*, vols I, II. Cambridge: Blackwell.

Sacks, H., Schegloff, E. A. and Jefferson, G. 1974. A simplest systematics for the organisation of turn-taking for conversation. *Language* 50 (4): 696–735.

Schwarz-Friesel, M. and Consten, M. 2011. Reference and anaphora. In W. Bublitz and N. R. Norrick (eds), *Foundations of Pragmatics*, vol. 1 of Handbooks of Pragmatics, pp. 347–372. Berlin: Mouton de Gruyter.

Scott, M. 1997. PC analysis of key words – and key key words. *System* 25 (2): 233–245.

Scott, M. and Tribble, C. 2006. *Textual Patterns: Key Words and Corpus Analysis in Language Education*. Amsterdam and New York: John Benjamins.

Smith, S. W., Hiromi, P. N., Andrews, S. and Jucker, A. H. 2005. Setting the stage: How speakers prepare listeners for the introduction of referents in dialogues and monologues. *Journal of Pragmatics* 37: 1865–1895.

Stenström, A. 1998. From sentence to discourse: *Cos* (*because*) in teenage talk. In A. H. Jucker and Y. Ziv (eds), *Discourse Markers: Descriptions and Theory*, pp. 127–146. Amsterdam and Philadelphia: John Benjamins.

Strauss, S. 2002. *This*, *that*, and *it* in spoken American English: A demonstrative system of gradient focus. *Language Sciences* 24: 131–152.

Tanz, C. 1980. *Studies in the Acquisition of Deictic Terms*. Cambridge: Cambridge University Press.

Wald, B. 1983. Referents and topic within and across discourse units: Observations from current vernacular English. In F. Klein-Andreu (ed.), *Discourse Perspectives on Syntax*, pp. 91–116. New York: Academic Press.

Walmsley, P. 2007. *XQuery*. Sebastopol, CA: O'Reilly.

Yule, G. 1996. *Pragmatics*. Oxford: Oxford University Press.