Notes to the text

Introduction

1. Most stories in the SCoSE apparently originated in conversations between academics based in Northern Illinois (cf. Norrick 2000).
2. A somewhat similar corpus project is reported on in Carruthers (2008). This project involves the collection, transcription, and encoding of 'new story-telling' in French, a type of narrative performed publicly to eclectic audiences.

1 Towards a working definition of conversational narrative

1. The inclusion of prayer among the subtypes of conversational narrative may be surprising. However, the clear parallel they share is that "both activities involve a quest for moral clarity and legitimacy" (Ochs & Capps 2001: 242). Note that this parallel even leads Ochs & Capps to "envisioning conversational narrative as an extension of prayer" (2001: 242).
2. While there may be little interactional *in the data*, as presented by Labov and associates, there is, of course, interaction *in the discourse*. As noted in the Introduction, narrators tend to obey the principle of recipient design, taking listener needs and expectations a priori into account. Seen thus, even Labovian stories are co-constructed; the type of co-construction is narrator-co-construction. What they lack decisively are traces of recipient-co-construction.
3. The functions alongside the basic 'vocalizing understanding' function include the function as 'continuer' (Schegloff 1982: 81), as 'convergence tokens' (O'Keeffe & Adolphs 2008: 85), and as 'supportive minimal responses' (Holmes & Stubbe 1997: 11). Further, investigating the temporal properties of minimal backchannels, Peters & Wong (forthcoming) found that certain high-frequency backchannels function as what could be called 'discontinuers,' signaling the listener's wish to take over the speaking turn.
4. As McCarthy (2003) observes, some backchannels in some contexts do gain a textual role contributing "to boundary phenomena such as openings and closings and topic structures" (McCarthy 2003: 35). This textual function, however, is restricted to non-minimal backchannels and it is as yet unclear whether such a textual role can also be observed for backchannels in conversational storytelling.
5. Note the necessary distinction between 'highpoint' (the most reportable event) and 'point' of the story (the reason why a story is told): while the most reportable

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

event is Carolyn's discovering that she'd have to face the intruder all on her own, the (likely) reason why the story is being told is the speakers' wish to empathize with Carolyn vis-à-vis the dangers inherent in her work.

6. Text labels in the NC, such as KC9-N2, refer to the source file in the BNC-C (the file KC9 in this case) and indicate whether the narrative is the first (N1) or second (N2) extracted from the BNC-C.
7. Norrick (2000: 3) maintains that an organizational principle competing with sequentiality is "repetition and formulaicity." His conversational data "exhibit storytellers organizing their performances around repetition and formulaicity as much as sequence" (2000: 3).
8. The mother's ability to guess that the boy chatted to Louise may, of course, be due to a number of contextual factors. One of these factors may be a successful reading of her daughter's use of 'foreshadowing,' a technique discussed in more depth in Section 1.7, where narrators insert into early stages of the narrative, details whose specific significance is revealed only later, when interpreted in the light of upcoming information. In the excerpt, the foreshadowing detail is the mention of Louise (*and he sat down next to Louise*), which, at the point in the discourse where it is inserted, seems marginal and hardly relevant. Only at the point where her daughter conveys her disappointment at the boy showing little interest in her, does the mother grasp the implicated significance of the foreshadowing mention of *Louise*.
9. Also, as Ervin-Tripp & Küntay (1997) observe, there is no explicit temporal sequence of events; the damage reported could occur all at the same time. However, given that the damage is the *consequence* of the earthquake, there is an "implicit temporal sequence" (Ervin-Tripp & Küntay 1997: 137).
10. A useful distinction of three types of material under the heading of 'orientiation' is proposed in Norrick (2000: 33): general frame (information about time and place); background information (additional information related to the setting, which may not be necessary for the point of the story), and narrow frame (information directly leading into the main action of the story).
11. Note that, in this quote, Labov, somewhat confusingly, equates the most reportable event with the story point (the reason why a story is told). More precisely, the most reportable event constitutes the story *high*point, or climax.
12. Obviously, this narrative macrostructure does not originate ex nihilo; among the cognitive processes involved may be the following. To begin with, consider Sacks's (1992: 237) notion of course-of-action organization, that is, the process of imposing a temporal sequence on the events to be reported. Bearing in mind that the succession of narrative clauses a and b is automatically interpreted as implying the a-then-b relation according to which a must have occurred before b occurred (cf. Labov & Waletzky 1967/1997: 24) the first, threefold, operation a would-be-storyteller must perform is (i) to select from all events that occurred in the told situation the story-relevant events by discarding events which have no bearing on the story, (ii) to sequentialize the relevant events, some of which may have occurred wholly or partly simultaneously, into a temporally coherent string of events, and, finally, (iii) following the clause-event match postulated by Labov & Waletzky's (1967/1997) (cf. Section 1.4), to encode events one by one in matching narrative clauses. That is, narrativization of experience entails a series of cognitive procedures designed to match the expectations that recipients

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

bring to narrative events. These procedures, it should be stressed, fundamentally re-organize experience imposing discoursal order on experiential disorder, thus revealing the "power of narrative to cast temporal order on the cacophony of daily life" (Ochs & Capps 2001: 45). Second, it is worth considering Labov's more recent notion of pre-construction. Labov (2006) postulates a recursive cognitive process whereby would-be narrators first decide on a reportable event and then proceed "backwards in time to locate events that are linked causally each to the following one" (Labov 2006: 37). While pre-construction may be theoretically intuitive, there is, to my knowledge, as yet little proof that indeed tellers pre-construct stories recursively.

13. It is worth noting the issue of credibility. Tellers need to employ ways to produce a story that is "recognizably correct, or at least possible" (Sacks 1992: 234). According to Sacks, credibility is ensured by use of "usualness or normalness measures" (Sacks 1992: 235), whereby scenes and events are described in terms of variances from the states the scenes are usually in and the ways that the events normally unfold. These usualness measures too obey the principle of recipient-design in that the relative unusualness of a scene as described by the teller can only be appreciated as such by the recipient on the basis of shared knowledge of what counts as 'usual.'

2 Data, methods, and tools

1. As noted in Rayson *et al.* (1997:135), there are 561 female speakers compared to 536 male speakers.
2. The numbers of words for the four social classes registered in the BNC-C are as follows: AB 81,7205, C1 782,234, C2 719,884, DE 451,485, UU 181,059.
3. At a more general level, Ochs & Capps's (2001) notion of narrative dimensions was useful too. This dimensional approach permits analyzing widely varied narratives within a unified framework. Ochs & Capps stipulate a set of five dimensions that "will be always relevant to a narrative, even if not elaborately manifest" (Ochs & Capps 2001: 19). The dimensions include Tellership (the extent and kind of involvement of conversational partners in the actual recounting of a narrative), Tellability (the extent to which the reported events are of interest and relevant to conversational partners), Embeddedness (the extent to which a narrative is detached from or embedded in the conversational context surrounding it), Linearity (the extent to which a narrative threads events into a unilinear time line and cause-effect progression) and Moral Stance (the way that a narrative encodes and perpetuates moral worldviews).
4. See www.oxygenxml.com.
5. Taken together, the proportions found for mediated stories – viz., for first-person (T1M), 3%, and, for third-person (T3M), 2% – match up well with the proportion reported in Blum-Kulka (1993) for adult stories mentioning characters from media such as books and films: while in children's stories this type of story accounted for 14%, in adult stories, it accounted for less than 4%.
6. Occasionally, exosituational orientation occurred *within* utterances. That is, it occurred only after the speaker had attended to some other discourse business. In that case, the utterance was broken in two parts at the point where the re-orientation

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

became noticeable; the resulting two parts were assigned to different components. Consider:

```
"Bruise the ginger" (embedLevel="ES" narrativeType="T10")
1     <u who="PS08Y">
        <s n="1672" sID="1672"/>
        <w c5="PNI" hw="none" pos="PRON">          None</w>
        <c c5="PUN">                               ?</c>
      </u>
    </seg>
→   <!– utterance split –>
    <seg Components="CNN">
      <seg Components="CNI">
2       <u who="PS08Y" Participation_roles="PNP">
        <s eID="1672"/>
        <s n="1673" sID="1673"/>
        <align with="KCoLCoWW"/>
        <w c5="PNP" hw="i" pos="PRON">             I</w>
        <w c5="VM0" hw="can" pos="VERB">           can</w>
        <w c5="VVI" hw="remember" pos="VERB">      remember</w>
        <w c5="AT0" hw="the" pos="ART">            the</w>
        <w c5="ORD" hw="first" pos="ADJ">          first</w>
        <w c5="NN1" hw="time" pos="SUBST">         time </w>
        <w c5="PNP" hw="i" pos="PRON">             I </w>
        <w c5="AV0" hw="ever" pos="ADV">           ever</w>
        <align with="KCoLCoWX"/>
        <s eID="1673"/>
      </u>
    </seg>
    (KCo-N1)
```

For space considerations only the (originally single) utterance that was split in two utterances, labeled 1 and 2, is given in the example. Inspection of the preceding context shows that participant PS08Y's *None?* in utterance 1 functions as a response to another participant who has just declined PS08Y's offer to have more ginger. In saying *I can remember the first time I ever* participant PS08Y performs a decisive shift in orientation leading the discourse from an offer-declination adjacency pair to the launch of a story, thus justifying the utterance split implemented in the text's annotation.

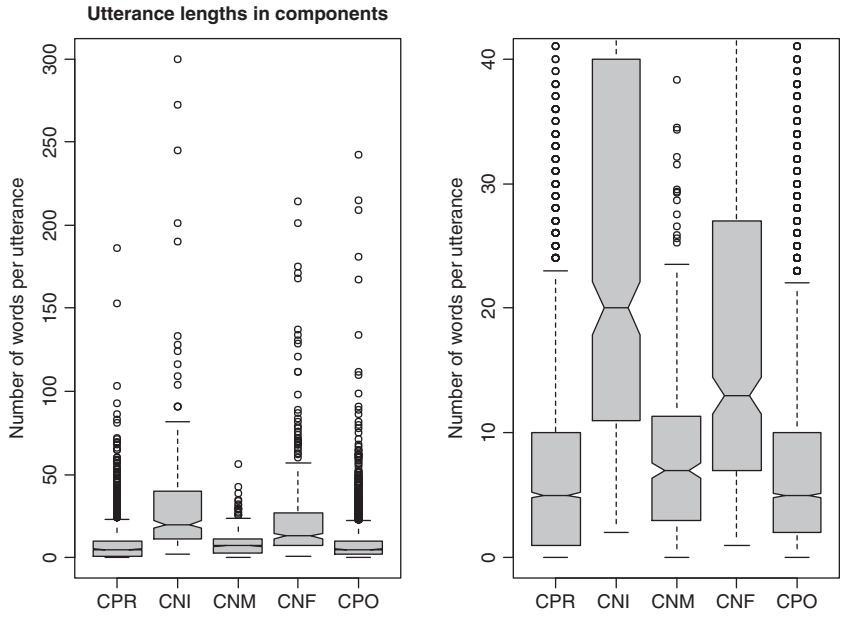
7. Interestingly, while the narrative component is larger than the surrounding non-narrative talk by roughly 8,000 words, it has strikingly fewer utterances. Taken together, CPR and CPO include 8,933 utterances, whereas CNN includes only 6,060 utterances. Utterances in narrative must hence be considerably longer than in general conversation.

   This is confirmed by a comparison of the mean numbers of words per utterance: the mean lengths of utterances in CPR (7.75) and CPO (8.07) are rather similar, yielding a combined mean of 7.91 words for the non-narrative components. Utterances in the narrative micro-components, by contrast, are much longer but also display much greater heterogeneity with regard to length. While CNI utterances have a mean length of 29.18 words (between three and four times

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

the mean lengths of utterances in CPR and CPO!), utterances in narrative-final position (CNF) are, on average, 22.55 words long (roughly three times as long as non-narrative utterances). The average length of utterances in narrative-medial position, however, drops to a mean of 10.1 words. Note that this drop is in no small part owed to the inclusion of recipient utterances; these are typically much shorter than narrator utterances, especially PRR-tagged utterances. Taken together, utterances in the CNN macro-components are, on average, 29.55 words long. In stories that are told largely single-handedly (without any feedback) and where the narrative constitutes one long uninterrupted utterance, tagged CNI-CNF, the mean utterance length is almost twice that high, namely 56 words per utterance.

That utterances in narrative are longer than utterances in 'normal' conversation is unsurprising and has long been noted (e.g., Ochs & Capps 2001). However, means only tell part of the story. What they cannot tell is how the data are dispersed and whether the observed differences are statistically significant. To approach these aspects, the lengths of utterances in the two non-narrative macro-components as well as the three narrative micro-components are graphed as boxplots:



In the figure, the boxplots showing the distributions of the lengths of utterances in the different textual components are presented in two different panels. While the left panel shows the distributions entirely (with the whiskers and all outliers included), the right panel presents the boxes in greater detail, thus permitting a clearer view of the boxes and the notches.

In the left panel, it can be seen that the dispersion in narrative-initial utterances (CNI) and narrative-final utterances (CNF) is much greater than in the non-

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

narrative components and also in narrative-medial utterances. This greater dispersion, and hence greater heterogeneity, is suggested, for CNI, not only by the larger range of outliers (indicated by the empty circles), but also, for CNI and CNF, by the greater distances of the whiskers (the dashed vertical lines) and the larger sizes of the boxes. This skewness shows up even more clearly in the asymmetry of the boxes for CNI and CNF, which becomes visible in the right panel. For these two boundary components demarcating the beginning and the end of the storytelling, the upper part of the boxes (above the median line) is much larger than the lower part (below the median line). Also, as shown in the right panel, the medians (indicated in the bold horizontal lines in the boxes) for CNI and CNF are much higher than for CPR and CPO, suggesting that narrative-initial and -final utterances tend to be much longer than non-narrative utterances. Note also that the median for CNI is still higher than the median for CNF, confirming that narrative-initial utterances are on average longer than narrative-final utterances. The medians for CPR and CPO, by contrast, are similarly low, while the median for CNM is still higher but clearly lower than for CNI and CNF. Further, only the notches (indicating the 95% confidence intervals) for CPR and CPO overlap, but not those for the three narrative components. This is prima facie evidence suggesting that the differences in means (observed above) and medians (observed in the figure) are significant (cf. Gries 2009a: 205). According to pairwise comparisons using Wilcoxon rank sum test (cf. Crawley 2007: 482), all but one pairing exhibit significantly different distributions: only the difference between CPR and CPO is insignificant ($p = 0.4$):

	CPR	CNI	CNM	CNF
CNI	< 2e-16	–	–	–
CNM	1.1e-06	< 2e-16	–	–
CNF	< 2e-16	2.0e-09	< 2e-16	–
CPO	0.4	2.0e-09	5.7e-06	< 2e-16

The reasons for the peaks in length in narrative-initial (CNI) and narrative-final (CNF) utterances cannot be explored in full detail here. A few, in part speculative, remarks may suffice. Given that the narrative-first utterance marks the change in discoursal orientation from the present situation to the anterior situation in which the events occurred, the dramatic increase in length that characterizes CNI may be associated with an increase in new vocabulary. Work, for example, by Youmans (1991) shows that breaks in the discourse (for example, transitions from one topic/section to another) may be correlated with vocabulary 'bursts', that is, sudden increases in discourse-new words. By contrast, the verbosity of narrative-final utterances is unlikely to indicate a similar influx of new vocabulary. Rather, it needs to be borne in mind that CNF, like CNI, overwhelmingly represents *narrator* utterances, whereas CNM comprises both narrator *and* recipient utterances. The length of CNF is therefore likely to a large part simply representative of the typical narrator utterance. This assumption is underscored by a comparison of the mean length of CNF, which is, as noted above, 22.55, and the overall mean length of narrator utterances (irrespective of narrator subrole), which is 21.07. According to a two-sided Wilcoxon rank sum test, this difference is not significant ($W = 742232.5$, p-value = 0.1999).

Second, I would argue that the increase in length from CPR to CNI and the decrease from CNF to CPO suggests that the segmentation of the texts in the NC

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

into the three major components pre-narrative talk (CPR), narrative (CNN), and post-narrative talk (CPO) indeed corresponds to segments in the discourse. To put it differently, I take the significant differences in utterance length in CPR and CNI as well as CNF and CPO respectively as evidence that the componential annotation in the NC reliably mirrors the shift in the generic framework from general conversation to narrative and from narrative back to general conversation. I also suspect that future *automatic* annotations of conversational narrative can use the two major points of changes in utterance length (from CPR to CNI and from CNF to CPO) as a means of identification of narrative boundaries.

Third, it is of interest to compare the utterance lengths obtained for the NC to the lengths obtained for the BNC-C. According to Rayson *et al.* (1997), women use, on average, 10.33 words per utterance, whereas men use 9.53 words per utterance (1997: 136). Rayson *et al.* do not present utterance length independent of sex but, given the figures for number of words and utterances per man and woman in the study, the overall mean length (independent of speaker sex) is 9.99 words per utterance. The mean utterance length obtained for the NC is strikingly similar: 9.97 words per utterance. As regards utterance length, the NC is thus a near-perfect reflection of the BNC-C. The near coincidence of the mean lengths can however be interpreted in another, wider, perspective. We can ask if the match in number of words per utterance between the two corpora indicates a match in the distribution of general conversation and conversational narrative in the two corpora. That is, considering that the NC consists of two roughly equal halves, one narrative (CNN) and one non-narrative (CPR and CPO combined), which are both drawn from conversation, does the near coincidence of the mean utterance lengths for the NC and the BNC-C indicate that, roughly, half of all conversation is conversational narrative? Estimates regarding the proportions of narrative in conversation abound. As noted in the Introduction, generally, very high frequencies are ascribed to stories in conversation. To my knowledge, these estimates are commonly not based on quantitative analysis and hence speculative. Admittedly, if I were to claim that storytelling makes up roughly 50% of all conversational talk such a claim would be not much less speculative given that the near coincidence of means of utterance length is just one, preliminary, piece of evidence and that many more factors may influence the generic make-up of conversation, a genre which is fraught with "genre-mixing and embedding" (McCarthy 1998: 31). However, the figures obtained for the NC do suggest that possibility and might serve as a starting point for more quantitative studies aiming to establish the exact proportion of storytelling in its host genre conversation.

8. To establish whether the differences in utterance length between participant types are significant, pairwise comparisons were performed using Wilcoxon rank sum tests. The results suggest that all possible pairings are significantly different:

	PRR	**PRC**	**PNU**	**PNS**	**PNP**
PRC	< 2e-16	–	–	–	–
PNU	< 2e-16	< 2e-16	–	–	–
PNS	< 2e-16	< 2e-16	0.00058	–	–
PNP	< 2e-16	< 2e-16	1.7e-11	< 2e-16	–
PNC	< 2e-16	0.01697	< 2e-16	< 2e-16	< 2e-16

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

9. Note, however, Barbieri (2007), who reports on the use of another 'new' quotative, *be all*, which is apparently not (yet) attested in many varieties outside North-America.
10. The tagset for BE *like* includes seven values: QBB = base form "be like," QBM = "'m/am like," QBZ = "'s/is like," QBR = "'re/are like," QBD = "was/were like," QBG = "being like," QBN = "been like."
11. Comparing this triplet to the quartet of SAY, GO, THINK, and BE *like* commonly seen as the major quotatives across most English varieties (see above), we notice the absence of BE *like*. It has a total frequency of a mere 9 tokens in the NC. Even if *like* (without preceding BE), which occurs 9 times too, is counted in as a variant of BE *like*, the resulting total frequency of 18 occurrences of (BE) *like* must be disappointing if compared to the increasingly high rates reported for this new quotative in much research. For example, in Tagliamonte & D'Arcy (2004), who look at quotative usage in Canadian youth, BE *like* was found to leave SAY, GO, and THINK far behind: BE *like* accounted for 58% of all quotatives in Tagliamonte & D'Arcy's corpus. Similar top rates have been reported for other varieties of English (for an overview of relevant research see Buchstaller 2008).

    Why is (BE) *like* so greatly underrepresented in the NC? The reason is straightforward: the recordings for the BNC-C, from which the NC is drawn, were made in the early 1990s, that is, at a time when the attested global spread of (BE) *like* from California, where it is commonly assumed to have originated (cf. Fairon & Singler 2006), to other varieties of English had only just reached the shores of Britain. In a number of studies based on conversational data from the first half of the 1990s, BE *like* was observed to have entered the language but still to trail far behind more traditional quotatives like SAY and also GO (cf. Miller & Weinert 1998, Andersen 2001, Rühlemann 2007). In today's usage in British conversational narrative, the situation may have dramatically changed: extrapolating, for example, from Tagliamonte & Hudson's (1999) study, in which BE *like*, THINK, and GO were equally represented (each accounting for 18%), BE *like* is highly likely, at least in narratives told by (female) adolescents, to challenge SAY's leading role seriously. Note also Buchstaller (2011), in whose NECTE2 corpus, which comprises Tyneside English from 2007–2009, BE *like* has increased to 21% overall (Buchstaller 2011: 73).
12. Given its metalinguistic focus, Representation of Use (MUU) may be somehow related to what is often called 'pure quotation' (cf. Brendel *et al.* 2011: 4ff.), as in "'Boston' is disyllabic" (Brendel *et al.* 2011: 4). The exact commonalities and differences between the two categories need yet to be explored though.
13. But see the many uncertainties that surround the notion of representativity and which, as noted in Leech (2007), make the quest for representativity resemble the quest for the holy grail – an ideal that will be never reached.
14. I'm grateful to members of the "Corp Ling with R" Google Group at http://groups.google.com/group/corpling-with-r for useful thoughts on the issue of population in corpus studies.
15. Nodes include not only 'elements' and 'attributes' but also 'root,' 'text,' 'namespace,' 'processing-instruction,' and 'comment.'
16. For this document to be fully complete as an XML document, also <?xml version="1.0" encoding="UTF-8"?>, the 'XML declaration' would have to be included.
17. To run XQuery, the (free) eXist XML database was used; cf. http://exist-db.org.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

3 How do narrators and recipients co-construct turntaking?

1. The following are extracts from examples (b) and (f) in Sacks *et al.* (1974: 702); the authors highlighted the single-word turns using arrows:

   (b) Jeanette: It was in the paper this morning.
   Estelle: It wa::s,
   → Jeanette: Yeah.

   (f) Bea: Met whom?
   Anna: Missiz Kelly.
   → Bea: Yes.

2. In this extract, the storytelling by the narrator is elicited by the recipient's question *Did you go to Andrew's this morning?* The question effectively accomplishes what was called 'exosituational orientation' (cf. Section 2.2.1), that is, the participants' re-orientation from the present situation to an anterior situation. The question is thus a crucial part of the storytelling process and is justifiably counted as the story's first turn.

   "Moving in" (Type: T10 / Embed Level: EC1)

   CPR
   (. . .)
   CNN
   CNI
   → S1 PRC Did you go to Andrew's this morning?
   S2 PNP No er, er no he came yeah [unclear]
   S1 PRC I mean did he take you this, I should say, did he take you this morning?
   S2 PNP And I, yes he came and so he says ^{QSZ} [^{MDD} I've got the keys to the flat] he said^{QSD}, [^{MDD}you don't want to come and look at it do you?] And I knew ^{QOO} [^{MII}he wanted me to].
   (. . .)
   (KD4-N2)

   In the NC, a total of 15 stories (3%) are initiated by a participant.

3. For readers interested in the details of, and possibly replicating, the simulation, I include here the complete code used in R:

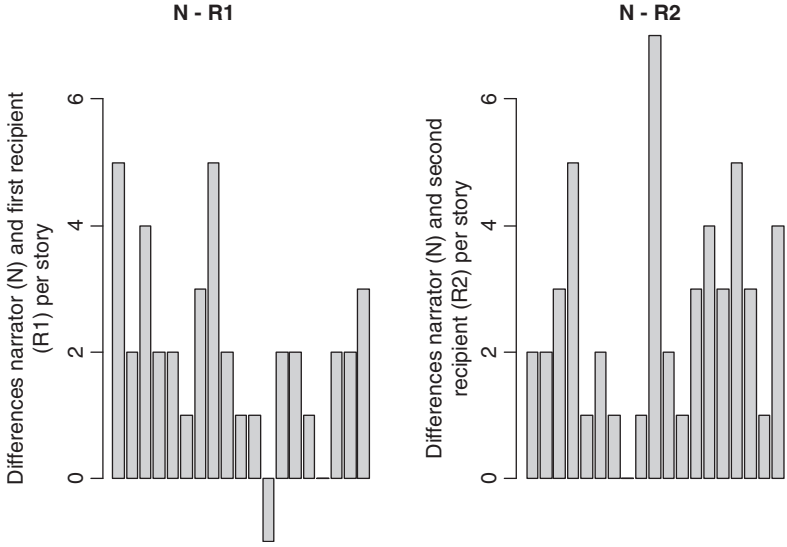
```
set.seed(123)
turns <- c("R1", "R2", "N")
proportions <- c()
for (j in 1:100) {
path <- c()
path[1] <- sample(turns,1,p=c(1/3,1/3,1/3)) # for N in first slot: path[1] <- "N"
for(i in 2:10000){
path[i] <- sample(turns[-which(turns==path[i-1])],1, p=c(1/2,1/2))
}
trigrams <- c()
```

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

```
for(i in 1:9998){
if(path[i]=="N" & path[i+2]=="N") trigrams[i] <- 1
if(path[i]!="N" | path[i+2]!="N") trigrams[i] <- 0
}
tab <- table(trigrams)/sum(table(trigrams))
proportions[j] <- tab[2]
}
mean(proportions)
```

Note that the restriction that the first turn slot is equally open to any of the three participants (and that hence the probability for them is 1/3 each) has *no influence* on the resulting proportion of N-notN-N trigrams. That is, even if N is set as defaulting to the first slot of the first trigram (a justifiable decision given that the vast majority of stories are started by the narrator him/herself; see the previous endnote), the proportion expected for N-notN-N trigrams is, according to the simulation, 17%. (To verify simply replace the sixth line of code with the code given after the "#".)

4. What is more, looking at the stories individually, the narrator gets virtually always more turns than each recipient. The barplots below show the *differences* in number of turns taken by the narrator and turns taken by the recipient who responds first (R1) (left panel) and, respectively, the recipient who responds second (R2) (right panel).



The number of narrator turns is smaller than the number of turns by one of the recipients (viz. R1) in one story, where the difference is hence negative (see the single downward bar in the left panel). The number of turns taken by the narrator

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

equals the number of turns taken by the recipient (R1 and, respectively, R2) in two stories, where the difference is hence zero. In all other cases, the narrator's turn share is clearly greater than any one recipient's.

5. The result was double-checked and fully confirmed by a bootstrap, the bootstrap's confidence interval (CI) spanning between 0.06 and 0.15.
6. According to Shapiro-Wilk tests for normality (cf. Gries 2009b: 150) the distributions of either variable violate normality (number_all_trigrams: p-value = 0.0005; proportion_N_notN_N: p-value = 0.017). The p-values are below the 0.05 threshold; therefore the H_o underlying the Shapiro-Wilk test, which assumes normal distribution, can be rejected.
7. The null probability does not impact the confidence interval (CI). The CI is the same for different null probabilities.
8. I'm grateful to Stefan Gries for suggesting the use of a permutation test for the present analysis.
9. In performing the bootstrap, the number of resamples was set to 10,000. That is, 10,000 different estimations of the proportion of N-notN-N trigrams out of all trigrams were generated. Based on these estimations and their differences a 95% CI was calculated. The most widely used CI (cf. Crawley 2007: 322) in bootstrapping is the BCa interval. It was found that the BCa CI obtained from 10,000 replicates ranges between 0.27 and 0.40. To compare, the CI obtained from the proportions test was 0.28 to 0.41.
10. To illustrate the difference between mean and median consider this set of five numbers 1, 2, 3, 4, 5. Both the mean and the median are 3. The numbers are distributed exactly symmetrically around this central value. If, however, the set of numbers is asymmetrical, as in the set 1, 2, 3, 4, 55, the mean changes to 13 while the median remains the central value, namely 3 (cf. Woods *et al.* 1986: 31).
11. The results for the non-narrative components CPR and CPO deserve a closer look. The medians and the IQRs for CPR and CPO reveal distinct tendencies: both fluctuation size and fluctuation homogeneity are markedly greater in the post-narrative component CPO than in the pre-narrative component CPR. They are clearly more similar to the median and the IQR obtained for CNN: TSF size and homogeneity behave in CPO more CNN-like than they do in CPR.To establish with greater confidence whether this 'spill effect' for TSF from CNN to CPO can be observed, not only for text KB9-N2, but *generally* for discourse segments following narratives in conversation is an interesting question worthy of further investigation in future research.
12. It may be felt that, compared to the medians and IQRs obtained for text KB9-N2, the reference values are too 'liberal.' There are two reasons to prefer these values to more rigid ones. First, the values obtained for CNN in the example text seem extreme; for example, an IQR of 0.15 is unlikely to be representative of many other texts. Second, one factor that is not controlled for in the present analysis and which inevitably leads to smaller medians and higher IQR values is overlap (simultaneous speech). Overlap is problematic for the present analysis in that TSF coefficients are calculated on the basis of the assignment of words to distinct <u>-elements (that is, utterances) which are invariably ordered successively and assigned to distinct speakers. As a result, in the case of overlapping speech this

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

practice can lead to the intrusion of 'fake' turns in the transcript and hence incorrect TSF coefficients. The issue is illustrated in the following extract; three columns are added to the text: the column labeled N gives the number of words per turn; next is the TSF coefficient (Coeff.), while the right hand column (Corr.) gives the corrected coefficients that would result if the overlap were taken into account

"Liverpool boys" (Type: T10 / Embed Level: EN)

	CNN CNI				N	Coeff.	Corr.
	1	S1	PNS	That's right. And all smoky. Oh I know. Well fo well there was some lads from Liverpool that was evacuated when well were evacuated. Oh! God! They were filthy!	30		
	2	S2	PRR	Mm.	1	0.94	0.94
	3	S1	PNS	I mean not just filthy i er bodily but they had filthy habits.	13	0.86	0.86
	4	S2	PRR	Mm.	1	0.86	0.86
→	5	S1	PNS	They were terrible! And in fact, I think they only la they didn't last a week. They were sent	20	0.90	0.91
	6	S2	PRR	Mm.	1	0.90	0.83
→	7	S1	PNS	back.	1	0.00	–
	8	S2	PRR	Yeah.	1	0.00	–
	CNF 9	S1	PNS	I'd like to have seen that.	7	0.75	0.75
				median:		0.65	0.82

(KBC-N2)

What the transcription presents as if they were two distinct utterances (the arrowed utterances 5 and 7) are obviously just a single turn to which the recipient's *Mm* in turn 6 is a response occurring simultaneously with *sent* (in turn 5) and *back* (in turn 7). The turn recorded as turn 7 thus constitutes a 'fake.' Its intrusion in the transcription has serious consequences for the TSF coefficients. Compare the medians: the median is 0.65 with the fake included as a turn in its own right but 0.82 if the overlap is corrected (that is, if the 'stranded' *back* is counted as the 21st word in utterance 5 and the recipient's *Mm* and *Yeah* are drawn together as constituting one turn). The difference in median is noticeable. Since overlapping speech is "very common in natural, informal conversation" (Crowdy 1994: 26), its cumulative effects on TSF coefficients are certainly not negligible. However, lacking evidence to suggest otherwise, overlap will occur equally often across the three textual components and the decreasing effect it has on TSF coefficients will thus be the same for them.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

13. To retrieve turn lengths for two-party narratives, this XQuery was used:

```
for $narr in //seg[@Components='CNN' and ancestor::div[starts-with
  (@narrativeType,'o')] and count(descendant::u)>7 and not(descendant::div
  [@embedLevel='EN']) and not(descendant::u[starts-with
  (@Participation_roles,'PX')])]
let $title := //div[@title and child::seg[.=$narr]]
for $u in //u[ancestor::seg[.=$narr]]
let $utt_length := count($u//w)
let $speakers := distinct-values($narr//u/@who)
where count($speakers[ends-with(.,'UNK')])=0 and count($speakers)=2
return
concat(string-join($title//div[@title]/string(@title),' '),';',$utt_length)
```

14. The number of coefficients is always the number of turns in a story minus 1. This is, as noted, because the coefficients are calculated on the basis of the lengths of two adjacent turns (turn bigrams).
15. It may be noted that the TSF medians for PRC (0.52) and PNC (0.50) are lower than the TSF medians for the non-narrative components CPR (0.60) and CPO (0.64). This is because in the present analysis only coefficients for turn-bigrams in which the narrator's utterance comes first and the recipient's response comes second were included; coefficients for turn pairs in which the recipient's response is first and the narrator's utterance is second were not taken into account.
16. The value of 0.3 as the median TSF – if turn size is *not* predetermined – is suggested by a simulation in R. In the code below, random turn lengths varying between 0 words (for non-verbal 'turns') and 300 words (the longest turn in the NC) are generated for a two-party conversation with 1,000 turns taken. The rounded median TSF of that conversation is 0.3:

```
conversation <- c()
for(turn in 1:1000){ # defines length of conversation (vz. 1000 turns)
conversation <- c(conversation, sample
(0:300,1)) # draws random turn lengths between 0 and 300 words
}
tsf <- c()
for(i in 1:(length(conversation)-1)){
tsf <- c(tsf,abs((conversation[i]-conversation[i+1])/(conversation[i]+
conversation[i+1]))) # calculates TSF
}
median(tsf)
```

4 Recipient design I: How do narrators mark quotation?

1. Leech & Short (1981) take care to qualify the assumed narratorial interference by use of the adjective 'apparent.' For, obviously, the narrator's power to interfere with the discourse presented is, in actual fact, always complete. Interference is thus not a gradient notion and proposing a 'cline' of narratorial interference is only defensible given the premise that certain discourse presentation types, particularly the direct

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

categories, implicate the narrator's *pretense* not to interfere with the discourse but to "leave the characters to talk entirely on their own" (Leech & Short 1981: 324).

2. The third internal speaker role is 'principal': "someone whose position is being established" (Goffman 1981: 144).
3. Indeed, Golato (2002) cites research suggesting that, like in other-quotation, whose non-verbatimness has been sufficiently demonstrated empirically (e.g., Wade & Clark 1993), self-quotation too is subject to selective depiction and thus far from faithful to the 'original' speech.
4. Note however that mimicry, which seems to occur in Direct and, particularly, Free Direct discourse presentation only (Mathis & Yule 1994, Stenström *et al.* 2002), may not only bring the reported utterance closer to the recipient but also signal 'quotative distance' (Stenström *et al.* 2002: 110) on the part of the mimicking speaker toward the mimicked person. Culpeper (2011), for example, discusses mimicry as an impoliteness strategy on the grounds that it can be seen as "a caricatured re-presentation" (Culpeper 2011: 161) or 'echo' of anterior discourse, "reflect[ing] the negative attitude of the echoer towards the echoed person" (Culpeper 2011: 165). In these cases, then, the immediacy of the report is relativized. See also Holt (2000), who stresses the 'duality' of dialog, which can "convey both the attitude of the reported speaker and, more implicitly, the attitude of the current speaker" (Holt 2000: 438).
5. As noted earlier, the quote (i.e., the reported clause) can be seen as the direct object of the reporting clause (Quirk *et al.* 1985: 1022–1023).
6. Because MSS presents (what the presenter has taken as) the utterance's illocutionary force, the recipient of the report is enabled to form an understanding of what was communicated and, given this understanding, to imagine locutions that convey the presented illocution. For example, in *[^{MSS} Nicola did n't believe me when I told her.]*, a recipient can easily construct potential locutions for Nicola's illocution: for example, *I don't believe you/ I can't believe you/ That's unbelievable* and so forth. That is, based on the reported illocution, the recipient can retrieve, not the actual, but possible locutions.
7. Brown & Levinson (1987) cite direct speech as a positive politeness technique operative in the service of "Strategy 3: Intensify interest to H [hearer]" (1987: 106).
8. In Clark & Gerrig's (1990) demonstration theory, the difference in use of reference is related to the difference in 'seriousness': non-direct reporting is 'serious' in the sense that it means what it encodes. For instance, to use Clark & Gerrig's (1990: 764) example (2), the Indirect report "she tells him that uh she wants to buy an ant" means the female referent informed a male referent that she wishes to purchase an ant. By contrast, a Direct speech report like "she says 'well I'd like to buy an ant'," Clark & Gerrig's (1990: 764) example (1), is 'nonserious' in the sense that the reporting speaker "isn't 'really or actually or literally' making a request" (Clark & Gerrig 1990: 770).
9. Clark & Gerrig (1990) report an experiment in which subjects were asked to memorize a brief scene from *Breakfast at Tiffany's* until they could recite it by heart. One group were asked to reproduce the scene 'as accurately as possible'; here, the verbatim reproduction rate was 99%. Another group were asked to 'simply tell the story of the scene' to a partner; there, the subjects produced verbatim quotations only 62% of the time (Clark & Gerrig 1990: 797); for more experiments providing support for demonstration theory, cf. Wade & Clark (1993).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

10. See Ameka's (1992) discussion of 'free' uses of *and*: he argues that *and*, when used as a stand-alone item, "usually carries a rising intonation signalling that it is an incomplete utterance" (1992: 105).
11. It has been noted in the literature that interjections are typical of direct quotation but untypical of (or even ungrammatical in) indirect or any other form of non-direct reporting (e.g., Mayes 1990). This observation is fully borne out by the data in the NC. The following are exhaustive lists of interjections occurring in (Free) Direct quotes and all other report units:

(Free) Direct units			All other report units		
N	ITJ	Freq.	N	ITJ	Freq.
1	oh	199	1	no	11
2	no	74	2	oh	10
3	yeah	49	3	yeah	8
4	ooh	34	4	yes	2
5	yes	29	5	aye	1
6	ha	15	6	hello	1
7	urgh	14	7	tt	1
8	blah	10			
9	hello	10			
10	ah	7			
11	hey	6			
12	mm	5			
13	aye	4			
14	ee	4			
15	aargh	3			
16	blimey	3			
17	cor	3			
18	dear	3			
19	eh	2			
20	hallo	2			
21	huh	2			
22	aah	1			
23	bye	1			
24	gee	1			
25	goodbye	1			
26	hah	1			
27	hi	1			
28	ho	1			
29	oops	1			
30	shush	1			
31	ugh	1			
32	wahey	1			
33	whoa	1			
34	whoop	1			
35	woh	1			

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

For a start, the two lists differ markedly in terms of length. The list for (Free) Direct comprises 35 types and 492 tokens, the list of interjections in all other modes contains merely 7 types and 34 tokens. Moreover, the two lists differ in terms of the kind of interjections included. The list for (Free) Direct is headed by *oh* with 199 occurrences and other conventionalized forms such as *no* and *yeah*. However, more characteristically it seems, there are a great number of much less conventionalized forms such as *ee*, *aargh*, *gee*, *wahey*, *whoa*, and so on. The typological range of interjections in non-direct units is much more restricted: the list mostly contains conventionalized forms, perhaps the only exception being *tt*. In sum, the contrast between the two frequency lists could not be greater: Are these differences significant? This question can be answered easily by use of the NC keyness analysis tool. The results are unambiguous: key POS analyses for ITJ in (Free) Direct compared (i) to (Free) Indirect and (ii) to all report types except the direct categories show the tag ITJ consistently ranked first, that is, as the POS tag most key; the p-values are far below 0.001 (note the loglikelihood values for keyness which are far above the critical value of 10.83):

Key POS tag in (Free) Direct vs (Free) Indirect ($p < 0.001$; critical value: 10.83)

Rank	Item	Freq. (Free) Direct	Freq. (Free) Indirect	Keyness (LL)
1	ITJ	492	9	130.042

Key POS tag in (Free) Direct vs all remaining report mode types ($p < 0.001$; critical value: 10.83)

Rank	Item	Freq. (Free) Direct	Freq. all other report types	Keyness (LL)
1	ITJ	492	34	216.241

12. The underlying formula to calculate positions is not without problems. For example, according to the formula, the second word in a three-word utterance has position 1/3 = 0.33, whereas the second word in a six-word utterance has position 1/6 = 0.17. That is, positions of words used *within* utterances (or any other unit such as quotes) are relative to, and hence dependent on, the length of the utterances. However, speakers hardly ever know in advance how many words long their utterances will be. Therefore, as far as within-utterance positions are concerned, it will be hard to make any large claims as to speakers being 'primed' to use words in particular positions. Fortunately, this caveat need not worry us in the present connection where the focus is on *first* position, which is invariably 0 whatever the length of the utterance.
13. To compute positions of interjections the following XQueries were used:
    for conversational utterances longer than 1 word (n>1):

```
for $u in //seg[starts-with(@Components,'CP')]//u[count(descendant::w)
  >1 and descendant::w[@c5='ITJ']]
let $n_w_utt := count($u//w)
for $itj in $u//w[@c5='ITJ']
let $n_w_prec_itj := count($itj/preceding::w[ancestor::u is $u])
let $pos := $n_w_prec_itj div $n_w_utt
order by $pos
return
concat($pos,' ',normalize-space(string-join($itj//w/text(),' ')))
```

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

for quotes longer than 1 word (n>1):

```
for $seg in //seg[starts-
   with(@Reporting_modes,'MD')][count(descendant::w)>1][descendant::w
   [@c5='ITJ']]
let $n_w_seg := count($seg//w)
for $itj in $seg//w[@c5='ITJ']
let $n_w_prec_itj := count($itj/preceding::w[parent::seg is $seg])
let $pos := $n_w_prec_itj div $n_w_seg
order by $pos
return
concat($pos,' ',normalize-space(string-join($itj//w/text(),' ')))
```

14. An obvious question arising at this point is the question *why* interjections are so strongly attracted to first position in utterances. For space limitations, pursuing the question in sufficient detail is far beyond my aims. A few notes must suffice here. First, it may be helpful to consider that interjections, being inserts (cf. Biber *et al.* 1999), do not enter into syntactic constructions with other words. Given their syntactic independence, interjections are an obvious choice for filling the first slot in talk (cf. Tao 2003). Further, drawing on Tao's (2003) observation that the definite article *the*, which is among "the most frequent initial items in written discourse" (Norrick 2009: 870), seems to avoid first position in utterances, Hoey suspects that the high use of interjections in first position serves "to avoid the appearance of (over-)definiteness that presumably fronting an utterance with *the* might be interpreted as implying" (Hoey, personal email communication). Finally, McCarthy (2003: 38) sees the utterance-first position "as the locus of choice where speakers frequently select items that contribute to the non-transactional stratum of talk." In this view, the utterance-first slot, as the juncture between two utterances, has an interpersonal focus, tying not only contributions but also contributors to one another.
15. Further analysis shows that the absence of *mm* from inclusion in quotes is not only restricted to occurrences in quote-first position but applies much more generally, viz. to other positions and also to quotes consisting of one word only. Its absence from quotation is thus virtually total.
16. Despite their low frequencies, the remaining eight items were examined statistically. It turned out that the differences in frequency of occurrence in first position in quotes as compared to utterances are insignificant for *yeah*, *no*, *yes*, *hello*, *ha*, and *blah*. The results for *ah* defy interpretation. While the p-value 0.0452 obtained from a Fisher's exact test is still smaller than 0.05, suggesting a significant association, the confidence interval poses a problem: not only does it contain 1 (which it should not if the result were indeed significant) but it also extends from 0.8 to infinity (a fact which is due to the 0 in the 2x2 table for *ah*). This is obviously far too large an interval to be confident about *ah*'s status in first position and to allow any definitive statement regarding significance. As regards the item *ooh*, this interjection is significantly differently distributed in first position in quotes and utterances: p-value = 0.003353, Cramer's V = 0.3695652. Consider:

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

I know but he was so funny, for a dog to be that perceptive, do you see what I mean? As soon as I came downstairs, [^{MDF} **Ooh** God, I shouldn't be in here] [laughing] and out he jumped ! (KD5-N2)

Despite its apparent significance, no claims will be made regarding *ooh* on the grounds that this item may not be seen as an independent interjection but rather merely a phonologically lengthened variant of *oh*.

17. A third possible interpretation is that "quoters use *oh* and *well* more frequently or in a wider range of ways because they believe that their use at the beginning of quoted speech makes the quotation sound authentic" (Hoey, personal email communication). This interpretation ties in well with the earlier proposed notion of the conversational utterance as the *model* for constructed dialog: in constructing the dialog, reporters design it in such a way that it maximally 'sounds' like real speech. The weakness to this view, however, is that it does not explain the *overuse* of *oh* and *well* in quotes. Only if the two items were used equally frequently in utterances and in quotes, as is the case for the interjection *yeah*, would this view have to be preferred over the two other interpretations discussed.
18. Yet another interpretation is worth considering: it could be argued that *well* and *oh*, because of their unusual frequency, cannot be seen as part of the quote but rather as discourse markers *outside* the quote, for example, as part of the reporting clause which introduces the quote (cf. James 1983; for a brief discussion of the issue see Jucker 1993). To support this view one could refer to the unusual frequency of the two items: they are overrepresented in comparison to the utterances they are meant to represent, hence they must constitute something else – something outside the quote. There are two weaknesses to this view. First, it builds on the verbatim assumption, which holds that quotation attempts to faithfully replicate locution, a view which is, as noted repeatedly, untenable for cognitive (speakers' memory limitations) and functional (speakers' selectiveness in demonstrating) reasons. Second, if it is assumed that *well* and *oh* are *over*represented and therefore not part of the quotes themselves it follows statistically that they are *under*represented within quotes. To illustrate, if *oh* is taken as a component of the quote, as in the analysis presented above, its frequency of occurrence in first position is 168, which accounts for 90% of all occurrences of *oh* within quotes (the proportion of occurrence of *oh* in *utterance*-first position was 74%, that is, significantly less). Conversely, if *oh* in quote-first position is *discounted* as not belonging to the quote, we are left with a total of 18 occurrences of *oh* within quotes. It turns out upon inspection that 5 occurrences are in what has become the new first position (which is, according to the 'old' count, exactly the second word in the quote). The percentage for *oh* in that position is merely 5/18*100 = 27%. In comparison to the proportion of 74% with which *oh* occurs in utterance-first position this is a dramatic *drop*, which is very highly significant: X-squared = 16.339, df = 1, p-value = 5.296e-05. An even more drastic discrepancy can be observed if the same recalculation is performed for *well*: X-squared = 58.2926, df = 1, p-value = 2.259e-14. The discrepancies would call for an explanation: why should speakers so blatantly underuse *oh* and *well* at the onset of quotes as compared to utterances? It is hard to imagine a plausible answer to this question. By underusing *oh* and *well*, which, in conversational utterances, are, as demonstrated, massively attracted to first position, narrators would fail to provide recipients with important clues as to the forthcoming quotation. Given the importance of recipient design governing

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

the production of narrative discourse, this non-provision is highly unlikely. Quote-first interjections are hence best seen as an integral part of the quote itself.

19. XPath to retrieve quotatives immediately preceding Direct quotes with *oh* in quote-first position:

```
//seg[@Quotatives and following-sibling::seg[1][@Reporting_
modes='MDD' and count(descendant::w)>1 and descendant::w[1]
[@hw='oh']]]
```

20. XQuery used to retrieve positions of quote-first *oh* and *well* relative to their enclosing narrative utterance:

```
for $u in //seg[@Components,'CNN']//u[count(descendant::w)>1 and
   count(descendant::seg[starts-with(@Reporting_modes,'MD')])=1 and
   count(descendant::seg[starts-with(@Reporting_modes,'MD')]//w)>1 and
   descendant::w[@c5='ITJ']]
let $n_w_utt := count($u//w)
let $quote := $u//seg[starts-with(@Reporting_modes,'MD')]
for $itj in $quote//w[1][@c5='ITJ']
let $n_w_prec_itj := count($itj/preceding::w[ancestor::u is $u])
let $pos := $n_w_prec_itj div $n_w_utt
order by $pos
return
concat($pos,' ',normalize-space(string-join($itj//w/text(),' ')))
```

21. The fact that silent pauses are far more numerous than *er* and *erm* is not the only reason why the analysis foregrounds silent pauses. There are two more reasons. First, the distinction between the filled pauses *er/erm* and silent pauses is not negligible in that *er* and *erm* sit somewhere at the boundary between words and non-words, while silent pauses are decidedly non-verbal. Secondly, in terms of annotation, *er* and *erm* are treated as <w> elements, while occurrences of <pause/> are given an XML tag of their own, which is a 'sibling' to <w> elements (cf. Section 2.3.2). This annotational difference has consequences for the calculation of positions: for example, because silent pauses are not counted as words, they can effectively obtain the value 1, while *er* and *erm* occurring within the quote can not, even if they are in quote-ultimate position. So, the positional values for *er/erm* and silent pauses would not be easily comparable.
22. In Rühlemann *et al.* (2011), long pauses were found to occur less frequently in narrative components (CNN) than in non-narrative components (CPR and CPO). They are thus likely to play a less central role in narrative than in non-narrative conversation, where they seem to be closely associated with extra-linguistic action.
23. XQueries used to retrieve positions of filled pauses:

```
for $seg in //seg[@Reporting_modes='MDF'][descendant::w[@hw='er']]
let $n_w_seg := count($seg//w)
for $pause in $seg//w[@hw='er']
let $n_w_prec_pause := count($pause/preceding::w[parent::seg is $seg])
```

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

```
let $pos := $n_w_prec_pause div $n_w_seg
order by $pos
return
concat($pos,' ',normalize-space(string-join($pause//w/text(),' ')))
```

Variables: 'er', 'erm'; 'MDF'; 'MDD.'

To retrieve positions of silent pauses:

```
for $seg in //seg[@Reporting_modes='MDF'][descendant::pause][count
  (descendant::w)>0]
let $n_w_seg := count($seg//w)
for $pause in $seg//pause
let $n_w_prec_pause := count($pause/preceding::w[ancestor::seg[.=$seg]])
let $pos := $n_w_prec_pause div $n_w_seg
order by $pos
return
concat($pos,' ',normalize-space(string-join($pause//w/text(),' ')))
```

Variables: 'MDF', 'MDD.'

24. A little more than 1,500 silent pauses seem to have been lost in the BNC's conversion from its earlier SGML format to XML (Hoffmann *et al.* 2008: 57), most of them in the context of multiword units such as *even if* and *at least* (Hoffmann & Evert 2007). However, given the 202,894 remaining silent pauses (Hoffmann, personal email communication), the loss may be considered negligible.
25. According to a Wilcoxon rank sum test, the distributions of the filled pause types *er* and *erm* were not significantly different (p-value = 0.3726), exhibiting exactly the same type of density curve. The two data sets for *er* and *erm* were therefore lumped together and treated as one data set.
26. For illustrative purposes, a subsample consisting of 10 positional values was generated, from which 4 resamples were drawn randomly:

Bootstrap resampling of a subsample of the pause data

Case	Original subsample	Resample 1	Resample 2	Resample 3	Resample 4
[1,]	0.00	0.76	0.00	0.00	0.00
[2,]	1.00	1.00	0.00	0.62	0.00
[3,]	0.64	0.62	0.00	0.00	0.00
[4,]	0.00	0.64	0.64	0.00	0.00
[5,]	0.00	0.00	0.64	1.00	0.64
[6,]	0.00	0.62	0.00	0.62	0.00
[7,]	1.00	0.00	0.00	0.00	0.00
[8,]	0.62	0.76	0.00	0.00	0.64
[9,]	0.00	0.00	1.00	1.00	0.00
[10,]	0.76	1.00	0.64	0.00	0.76
mean	0.45	0.54	0.29	0.32	0.21
sd	0.44	0.4	0.39	0.44	0.33

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

As can be seen in the table, the value 0.76 (Case [10,] in the original subsample) is not included in Resamples 2 and 3, whereas the value 0.62 (Case [8,] in the subsample) is found twice in Resamples 1 and 3. Given these different values in the subsample and the resamples it is not surprising that both the means and the standard deviations are different too – in fact, they are quite different: the mean of Resample 1, for example, is 0.54, while the mean for Resample 4 is only 0.21. Similar differences can be seen for the standard deviations. Obviously, the differences are so stark due to the extremely small sample size of 10; they will decrease with increasing sample size.

27. On the other hand, the peak for pauses in final positions is less marked than the peak for initial positions. This might be taken as evidence in defense of the assumption that pauses primarily serve to buy the narrator planning time: it might be seen as reflecting the lesser processing costs of re-instantiating one's own reference and voice system.

5 Recipient design II: How do narrators use discourse presentation for dramatization?

1. XQuery to retrieve sequences of n report units within utterances:

```
for $u in //u[ancestor::seg[@Components='CNN'] and
   count(descendant::seg[@Reporting_modes])=3]
let $rep := count($u//seg[@Reporting_modes])
return
concat($rep,';',normalize-space(string-
join($u//seg[@Reporting_modes]/string(@Reporting_modes),'')))
```

Variables: n = 3, 4, . . . , 8.

2. In order to get an idea of the distribution of the ordinal values within each sequence we could plot the column totals, that is, the totals of the values for each position. However, the huge discrepancies in frequencies of the sequences produce hugely different column totals which it would be hard to compare visually. The column totals will be used later though in performing the tests. A measure of central tendency could be used instead. Since we are dealing with ordinal values, the median – "the score which divides each set of scores into two nearly equal halves" (Woods *et al.* 1986: 28) – would be a natural candidate (Gries 2009b: 107). However, the median turns out to be almost always the same: namely 8, the value for MDD, which is, as we know, by far the most frequent report type and therefore by far the most likely 'middle value.' To summarize the data for visual exploration, the ECDFs were therefore preferred (thanks are due to Stefan Th. Gries for suggesting them on the "Corp Ling with R" Google Group (http://groups.google.com/group/corpling-with-r)).
3. While the computation of the observed column totals is obvious, the value corresponding to 100% is calculated by multiplying the highest value afforded in the immediacy cline, namely 9, by the total number of occurrences of each sequence type (n=3, n=4, etc.). To exemplify, as shown in Table 5.5, there are 125 utterances with n=3 sequences. The 100% value is hence 125*9 = 1125. The test for trends in proportions juxtaposes this maximal value with the column totals, which,

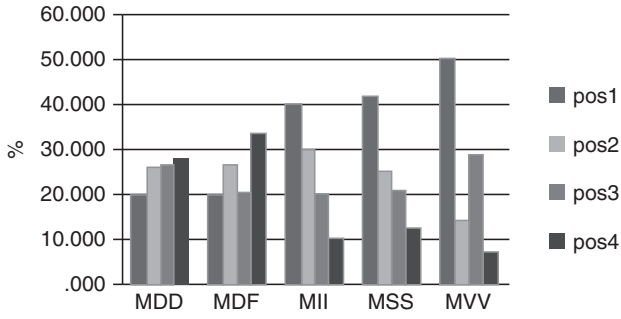
Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

in the case of n=3-sequences, are 827 (in pos1), 901 (in pos2), and 900 (in pos3). The column totals and maximal values ('max.' in Table 5.5) are the following:

Table 5.5

	pos1	pos2	pos3	pos4	pos5	pos6	pos7	pos8	max
n = 3	827	901	900						1125
n = 4	424	470	457	485					576
n = 5	190	222	219	233	208				270
n = 6	124	173	161	164	171	161			198
n = 7	71	82	96	94	86	91	93		108
n = 8	77	94	72	75	85	89	94	91	108

4. Another methodology would be to single out data sets for manual analysis. Take for example the data set for n = 4. This data set can be processed in the following way. First, frequency lists for all report unit types occurring in the set as a whole are generated. Second, frequency lists for all report unit types per position (that is, per column) are computed. Finally, the percentage values are computed for the report unit types per position as proportions of the total frequencies. To illustrate, the tag MII (Indirect) occurred 20 times in n=4-sequences as a whole while it occurred 8 times in pos1 (first position). Hence, its proportion of occurrence in that position is 8/20*100 = 40%. The figure below represents these proportions for all tags with total frequencies greater than 10.



The evidence presented in the figure is striking: only the proportions of occurrence of MDD and MDF, the two report unit types located at the very top of the immediacy cline, exhibit a rise from pos1 to pos4; all other types, including MII (Indirect), MSS (Speech Act), and MVV (Voice), see a clear fall from pos1 to pos4. The obvious drawback to this methodology is its costliness: performing this type of manual analysis for all *n*s as well as carrying out tests for trends in proportions for all sequences is highly labor intensive.

5. To illustrate, there were 30 utterances with n = 5; the maximum column total is hence 30*1 = 30. The observed column totals were: 18, 21, 23, 25, and 18. Thus, to perform the test, two vectors were created: one containing the maximally possible total, given five times (for each position), 30, 30, 30, 30, 30, and one containing the observed totals 18, 21, 23, 25, 18.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

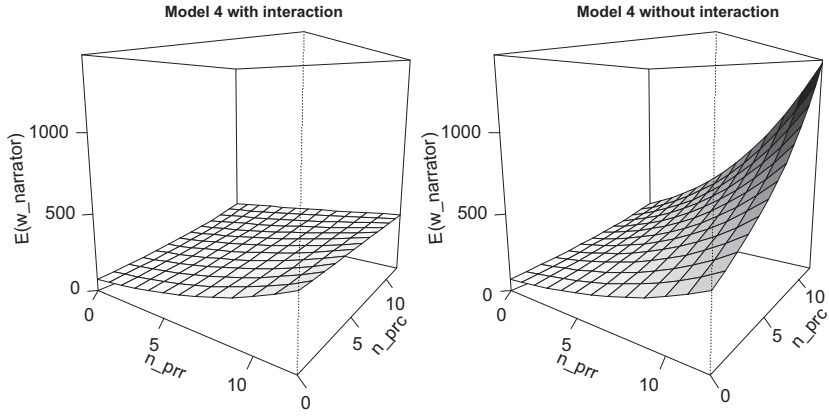
6 How do recipients co-author stories?

1. The boundaries between the word classes are anything but clear-cut. As Stubbs (2002: 40) notes, modal verbs, for example "express obligation, permission, and ability, and therefore convey content"; another problematic member of the class of function words are pronouns, which "can have extralinguistic reference" (Stubbs 2002: 40).
2. There are altogether 179 non-verbal PRR utterances (of which the overwhelming majority consist of laughter tokens), accounting for 14% of all 1,264 PRR utterances recorded in the NC.
3. XQuery used to retrieve data for subset 1:

```
for $cnn in //seg[@Components='CNN' and count(descendant::u
   [@Participation_roles='PRR'])>0
   and count(descendant::u[@Participation_roles='PRC'])=0]
let $w_narrator := count($cnn//u[starts-with(@Participation_roles,'PN')]//w)
let $n_prr := count($cnn//u[@Participation_roles='PRR'])
return
concat($w_narrator,'',$n_prr)
```

This XQuery instructs the processor to loop over stories (CNN) which meet two conditions. First, the number of PRR is greater than 0; this constraint excludes stories told by PNU (which by definition does not receive any recipient feedback). Second, the number of PRC contributions is set to 0; this excludes stories in which both PRR and PRC contributions occur. The query outputs two types of data: number of words in narrator utterances per narrative ($w_narrator) and number n of PRR contributions per narrative ($n_prr).

4. In the present analysis, the NB model is preferred over the Poisson model in that the latter requires equidispersion – that is, "in the Poisson model, the variance of the random variable is constrained to equal the mean" (Greene 2008: 585) – a requirement not met here: the variances are greater than the means.
5. The interaction effect is brought out more clearly if the model with the interaction is compared to what the model would look like without the interaction:



Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

The figure makes immediately obvious that if there were no interaction effect between n_prr and n_prc, the growth in w_narrator would be much more dramatic, as shown in the steeply rising plane in the plot on the right, than is the case when the interaction is taken into account. Its effect can be seen in the much flatter plane in the plot on the left. Note, however, that, for the sake of comparability, the scale of the plot for the model with the interaction has been changed to match the scale for the plot without the interaction. Note also that to more correctly illustrate the interaction effect it would be necessary to juxtapose Model 4 and Model 3 (for PRR + PRC without interaction); in that comparison, however, the interaction effect would not become salient enough.

6. In van Dijk & Kintsch's terms, the 'textbase,' which "represents the meaning of a text in all its detail" (van Dijk & Kintsch 1983: 52), is expanded by the recipient.
7. Fox & Thompson (2010: 136) refer to this type of *wh*-question as 'telling question' (which, on Fox & Thompson's account, contrasts with 'specifying question').
8. It will be noticed that in the narrative in (6.10), the first three consecutive turns were marked as Primary Narrator (PNP) turns. This annotation is correct due to the fact that the narrative is the second (embedLevel EC2) in a story round and that it overlaps with the first (EC1): what appears as the second PNP turn in EC2 is in fact the concluding turn in EC1.
9. This is because instances of discourse presentation were marked up using <seg> elements not only within <u>-elements (utterances) that have attribute values for narrator roles PNP, PNC, etc. but also within <u>-elements with values for recipient roles. So queries for discourse presentation segments contained within <u>-elements that have the attribute value PRC can be conducted and exhaustive quantitative analysis is possible.
10. The total number of PRC report units in the NC in its present version is 120. Upon close inspection, however, for four occurrences of reporting the annotation was found erroneous.
11. An example of Free Indirect discourse presentation by Co-constructive Recipient not included in the NC is given in this excerpt from the BNC-C. The main narrator, speaker PS18L, is relating a session at a fortune teller's. When she describes her skeptical thoughts about the weird looking room, a co-participant, speaker PS18K, produces the thought presentation *What are [you] letting yourself in for?*. This presentation is indirect: note the participant's switch into the primary narrator's deictic system evidenced by the reflexive pronoun *yourself*:

PS18L >: Exactly! We went into this room and there was this music playing like on the erm . . . like John
PS18E >: Indian music.
PS18L >: would play on the
PS18E >: Oh no.
PS18L >: on erm
PS18K >: Like in fantasy
PS18L >: Yeah.
PS18E >: John's sort of thing.
PS18L >: You know water [. . .] . . . a blue light bulb in the middle of the room
PS18E >: [laugh]
PS18L >: and I'm thinking, Oh!

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

PS18E >: [laugh]
PS18L >: Come on!
PS18K >: What are [yoo] letting yourself in for?
PS18L >: You know. But it was just . . . I mean, the things that they said . . . honestly . . . weird.
(KBU)

12. The p-values reported in this paragraph were obtained from chi-squared tests.
13. Note that epistemic modality (the level of reliability of the speaker's knowledge) is *included* in Clift's (2006) notion of evidentiality, whereas it is *excluded*, for example, in Aikhenvald's (2004: 3) definition of evidentiality as "a linguistic category whose primary meaning is source of information." Although the more inclusive notion is prone to "conceptual and terminological confusion" (Aikhenvald 2004: 5), it has a wide currency and will therefore be adopted in the present analysis.
14. In this case, then, the sequential placement of the recipient quote *after* the narrator's implicit assessment would make an interesting contrast with the sequential order observed by Holt (2000), which is the converse in that, there, tellers first produce a quote to which recipients then offer an assessment (Holt 2000: 450).
15. Extracts (6.23) and (6.24) are counterexamples in that here the recipient mimics an utterance by a speaker who is not co-present in the telling situation. My sense is that, in assuming the character's role, the recipient does not affiliate with the character but rather takes the dissociative attitude of mockery. The quote thus serves, not identification, but, via ridicule, dissociation.
16. 'Laughter', here, refers to both (i) between-speech laughter, coded as [laugh] and tagged, in the BNC and the NC, as a <vocal>element, and (ii) within-speech laughter, coded as laughing voice quality and tagged as a <shift>element. The two types have a combined frequency of 82 occurrences in 26 narratives containing recipient dialog, whereas, in the 505 narratives without recipient dialog, there are 524 instances of laughter. According to Fisher's exact test, this difference is very highly significant: p-value = 4.8e-07 (odds ratio = 0.33; CI: 0.20 - 0.53).
17. To retrieve positional data for quotes by recipients (PR*) and, respectively, narrators (PN*):

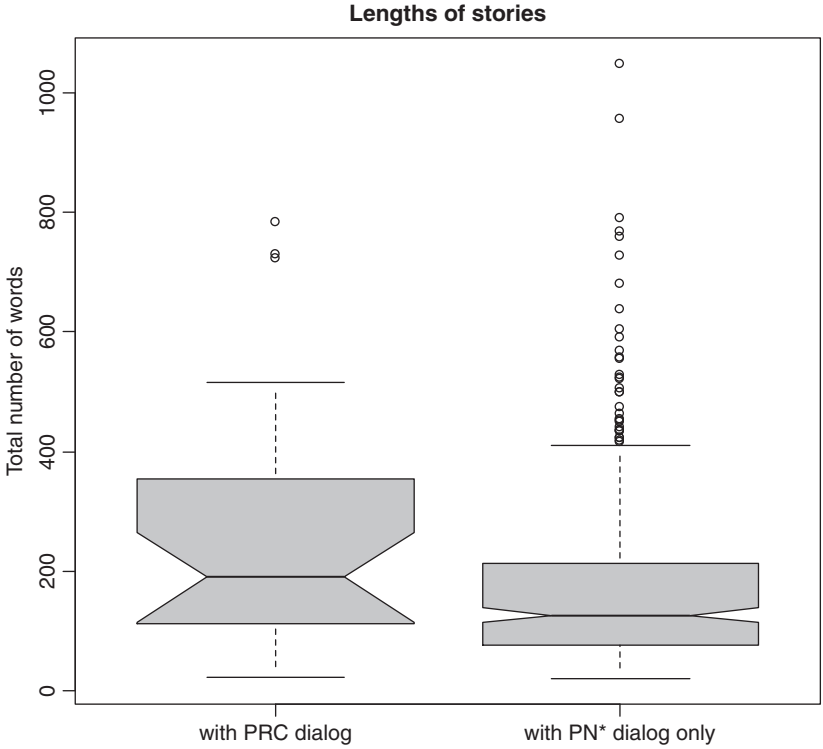
```
for $seg in //seg[@Components='CNN' and descendant::seg[starts-
   with(@Reporting_modes,'MD')] and
   not(ancestor::div[@embedLevel='EN'])]
let $n_quotes := count($seg//seg[starts-with(@Reporting_modes,'MD')])
let $n_words_all := count($seg//w)
let $n_words_inside_quotes := count($seg//w[ancestor::seg[starts-
   with(@Reporting_modes,'MD')]])
let $n_words_outside_quotes := $n_words_all - $n_words_inside_quotes
let $title := //div[@title and child::seg[.=$seg]]
for $targetquote in $seg//seg[starts-with(@Reporting_modes,'MD')]
   [ancestor::u[starts-with(@Participation_roles,'PR')]]
let $n_words_preceding_targetquote :=
count($targetquote/preceding::w[ancestor::u[ancestor::seg[.=$seg]]])
let $n_quotes_preceding_targetquote := count($targetquote/preceding::seg
   [starts-
   with(@Reporting_modes,'MD') and ancestor::u[ancestor::seg[.=$seg]]])
```

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

```
let $n_words_in_quotes_preceding_targetquote := count($targetquote/
    preceding::seg[starts-
    with(@Reporting_modes,'MD')and ancestor::u[ancestor::seg[.=$seg]]]//w)
let $pos := ($n_words_preceding_targetquote - $n_words_in_quotes_
    preceding_targetquote) div $n_words_outside_quotes
return
concat(string-join($title//div[@title]/string(@title),''),';',$n_words_all,';',
$n_quotes,';',$n_words_preceding_targetquote,';',$n_quotes_preceding_
targetquote,';',$n_words_in_quotes_preceding_targetquote,';',$n_words_
outside_quotes,';',$n_words_inside_quotes,';',$pos)
```

Variables: 'PR', 'PN'

18. One reason why the new formula impacts on positions of quotes in narrator talk but has no such impact in recipient talk is length of story: in long stories, the number of words within quotes is just a fraction of the total number of words and hence negligible. So it was examined whether stories in which recipient dialog occurs are longer than stories in which only narrators produce dialog – this is indeed the case (according to a Wilcoxon rank sum test, p-value = 0.006146):



Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

As indicated in the boxplot, the median length of stories where only narrators (PN*) produce dialog is 126 words, whereas the median length of stories where PRC dialog occurs is 190 words. The greater length of stories with PRC dialog not only explains, to an extent, why the new formula does not produce significantly different positions for PRC dialog; it also raises interesting questions as to whether extended length of story might be one factor making the use of recipient dialog more likely – a hypothesis worth testing in future research.

19. To avoid ties (i.e., identical values), a little 'noise' was added with the function jitter() (cf. Baayen 2008: 74).
20. As on previous occasions, I'm indebted to a fruitful discussion on Stefan Gries's "Corp Ling with R" Google Group of the issues involved in testing for sameness and accepting or rejecting insignificant results.
21. The narrative given in example (6.23) is a corrected version. In the present version of the NC, the same stretch of discourse is, erroneously, seen as part of a longer narrative:

"Wurgh!" (Type: T10/Embed Level: EC1)
CNN
CNI
S1 PNP Like when Lisa dropped one of the fridges on the floor, everybody used to go [MDD Wurgh! Urgh! Hey eh!]
S2 PNC [MDF Wurgh urgh!]
S3 PRC You did, didn't you? I know I was [. . .] when I worked in fruit packing, it's like someone drops an orange, [MDF Urgh urgh urgh]
S1 PRR [laugh]
(KPR-N2)

This annotation however misses two points. First, speaker S2's quote *[MDF Wurgh urgh!]* does not presuppose any privileged witness knowledge, as is required for the assignment of the tag as Ratified Co-narrator (PNC); rather, the response is clearly a Co-constructive Recipient's response and would thus have to be tagged PRC. Second, the discourse produced by speaker S3 is a follow-up story rather than an extension of speaker S1's story. This can be inferred from the new exo-situational orientation, evidenced in the clause *when I worked in fruit packing*, which re-orients the discourse to an entirely different situation, whose main coordinates – actor, place, and time – have changed. Instead of one story we have here two stories. The mistaken annotation will be corrected in an updated version of the NC.

7 Conclusions and implications

1. There is, indeed, very strong evidence that the N-notN-N pattern is operational not only in three-party but generally in multi-party narrative as a whole (cf Rühlemann & Gries forthcoming)
2. The following is a three-party narrative whose TSF pattern is both sizable and homogeneous (cf. the median and the IQR presented at the bottom of the excerpt):

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013

"Stew" (Type: T10/Embed Level: EC2)
CPR
(. . .)
CNN
CNI

			N	Coeff.
S2	PNP	Oh yeah, what you reckon Joy I did it cos we went to see his sister yesterday, cos he's only got one sister so we go and see her, see her regular and, she's not all that good is she in health?	44	
S1	PNC	No	1	0.96
S2	PNP	So she said [QSD [MII she's no car to come over here]	11	0.83
S3	PRR	Mm	1	0.83
S2	PNP	we go over there and erm so I do, what I did I put that meat out to thaw the night before so it was thawed	26	0.93
S3	PRR	Mhm	1	0.93
S2	PNP	so I thought [QTD [MDD well I'd put that in with some onion]	12	0.85
S3	PRR	Mm	1	0.85
S2	PNP	and, so I did, you know, [MDF and, and pearl barley in it]	12	0.85
S3	PRR	Yeah	1	0.85
S2	PNP	and then er I thought [QTD [MDD oh I might as well put some veg in,] you know, so I put some veg in, so when we, I says [QSZ [MDD oh I'll do this, I'll do this stew,] you know, [MDD nearly to finish it like, and erm when we come in we'll have a meal ready for us,] it were right nice coming into it meal ready	67	0.97
S3	PRR	Yeah, yeah	2	0.94
CNF				
S2	PNP	weren't it love?	4	0.33
S1	PNC	Yeah	1	0.66
			median:	0.85
			IQR:	0.1

(KB2–N2)

3. According to a Wilcoxon rank sum test, p-value = 0.003401.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 14 Nov 2026 at 16:42:22, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9781139026987.013