# 5
# Disfluencies

This chapter deals with disfluencies, in particular, with disfluency pauses. Fluency is one of the features of CDS that has been extensively discussed in the literature. Kidd et al. (2011) describe fluency as one of the 'hallmarks' of CDS. As discussed in Section 1.2, CDS is typically more fluent than ADS (Snow 1996; Broen 1972). In Section 5.1 I will present the types of disfluencies that have already been investigated in studies of CDS in other languages. In Section 5.2, I introduce the types of disfluencies evident in the Qaqet pear corpus, and link them to a model proposed by Clark and Wasow (1998). Unfilled pauses, called disfluency pauses in the following text, are the most common form of disfluency in the data. In Section 3.2.4, I explain their subdivision into hesitation pauses and self-interruptions. In Section 5.3, I will compare disfluency pauses in ADS and CDS in Qaqet, and provide an explanation for the results. Finally, in Section 5.4, in a subcorpus of the pear stories, the positions of hesitations in the sentence will be analysed to investigate whether they occur at positions that may be helpful for the listener, for example, by announcing difficult-to-process linguistic material (Kidd et al. 2011).

## 5.1 Previous research on disfluencies in CDS

The design of the present study was shaped by several studies that have found fewer disfluencies in CDS than ADS. For example, Broen (1972) showed that the disfluency rates in English differ significantly in speech to adults, speech to children over 45 months, and speech to 18–26-month-old children in a free play condition. The younger the listener, the fewer

the disfluencies. For a storytelling condition, the disfluency rates for speech to older and to younger children were comparably low, and both were lower than speech to adults. Thus, as the current study used a narrative condition, it was expected that the rates for speech directed at children over 45 months and under 45 months would both be significantly lower than those in adult-directed speech. Similar evidence, although not for narratives, was provided by Kidd et al. (2011), who investigated disfluencies in CDS data from the CHILDES database. They found a lower rate of hesitation particles than reported in the literature for adult-directed English. Nilsson Björkenstam et al. (2013) showed that there are significantly more disfluencies in ADS than in CDS in their longitudinal recordings of Swedish parent-child interactions with children from six to 33 months. All three studies also stated that the rate of disfluencies increases with the age of the child.

In Section 1.2 it was discussed why a low rate of disfluencies is not a conscious effort to make comprehension easier for the child. The discussion demonstrated that it is improbable that CDS is more fluent *because* the added fluency supports hearers' processing. Instead, it may be the lower complexity, as measured by MLU, that is responsible for the fluency of CDS. Rispoli and Hadley (2001) showed that, for children's productions, sentences with disfluencies tend to be longer and more complex than fluent sentences. Accordingly, in Section 5.3, the length of intonation units with and without hesitations will be compared.

I also presented evidence that hesitations are not necessarily detrimental to comprehension (see Section 1.2). Rather, hesitations occurring before difficult material may even aid listeners by reducing their processing load by announcing complications (Kidd et al. 2011) or relevant boundaries (Snow 1972). The results of those studies suggest that even young children can make use of them and modify their expectations towards following referents accordingly (Kidd et al. 2011). These findings were confirmed by Owens and Graham (2016) for three-year-old children, but not for two year olds. Orena and White (2015) found that children's evaluation of speaker knowledge influences their expectations towards the role of filled pauses in speech. Recently, Owens et al. (2018) presented opposite results, suggesting that children are not sensitive to the pauses, but rather to the semantics of the words, whereas the sensitivity to disfluencies develops later. The studies whose results have been reported above have defined disfluency in very different ways. Broen (1972: 10), for example, counted 'repeated or interjected sounds, words, or phrases'

among the disfluencies while Kidd et al. (2011) considered only filled pauses. These are also addressed briefly in this chapter, but the central disfluency addressed are unfilled disfluency pauses, that is, hesitation pauses and self-interruptions as defined in Section 3.2.4. All disfluencies have been annotated following the CHAT conventions (MacWhinney 2009) on a separate tier in ELAN. The next section will illustrate the operationalisation of disfluency I developed for the present study with reference to a model proposed by Clark and Wasow (1998).

## 5.2 A model of disfluency pauses with reference to Qaqet

As already described in Section 3.2.4, in Qaqet, continuation marked intonation units, that usually show some kind of processing trouble, carry final level pitch, and the constituent before the pause is glottalised (Hellwig 2019: 140). During annotation it became clear that for both CDS and ADS, nearly all disfluency pauses occur with preposed function words, as shown in (38). This is true both for hesitation pauses and self-interruptions.

(38) *tatrama.. amagama*

| | | | |
|---|---|---|---|
| ta=tat | ama | ama=gam-a | […] |
| 3PL.SBJ=take/pick_up | ART | ART=seed/fruit-DIST | |

'they pick the.. the fruits […]' (PearARLA 079)

In example (38) the article *ama* is 'realized with level pitch and a final glottal stop as [amaP] […], signalling hesitation and the continuation of the utterance' (Hellwig 2019: 140). A pause follows, after which, in the case of hesitation pauses, the article is repeated, this time attached to the noun that is its lexical host. In the case of self-interruptions, of course, the utterance is abandoned after the pause. Clark and Wasow (1998) propose a four-step model for disfluency pauses that explains this process (see Table 5.1).

Instead of simply pausing when they have problems with a constituent, people who feel pressure to talk may choose to 'commit to a constituent by producing its first word or words […] as early as possible' (Clark & Wasow 1998: 235) (Stage I in Table 5.1).

**Table 5.1: Four stages in repeating a word (table from Clark & Wasow 1998: 235), Qaqet examples added.**

| Stage | Speaker S's action | Spoken example from Qaqet |
|---|---|---|
| I. Initial commitment | S commits to a constituent | *ama* 'the' |
| II. Suspension of speech | S stops vocalising | { |
| III. Hiatus | S deals with potential delay | level pitch + final glottalisation |
| IV. Restart of constituent | S restarts the constituent, restoring continuity to it | } *amagataqi* 'the basket' |

The insertion of function words before a hesitation pause can be explained by their low activation threshold (Levelt 1989: 203). Either the frequency or the predictability of a word is responsible for the low threshold. Clark and Wasow assume that people are 'pressed by a temporal imperative':

> If they delay too long, they may be heard as opting out, as confused or distracted, as uncertain about what they want to say, or as having nothing immediately to contribute. They can forestall these attributions by producing the first word of the next constituent (even if prematurely) to show that they are engaged in planning the constituent. (Clark & Wasow 1998: 238)

It is reasonable to assume that the experimental situation in the present study, telling the pear stories, made the participants feel a certain pressure to speak. They were in an unknown situation, being asked to retell a story and were even being recorded doing so. This may explain the large number of preposed function words uttered before hesitations found in the data. Still, impressionistically, this pattern also seems to prevail in the longitudinal recordings of our project (Hellwig personal communication).

Stage II in Table 5.1, the stop in vocalising, may be caused by various factors. At any level of speech planning, difficulties may arise so that the speaker needs additional processing time. These difficulties might include 'problems in segment retrieval or in implementing the articulation of the next syllable, replanning of the overall message or the grammatical structure, on-the-fly changes in lemma selection' (Himmelmann 2014: 950) or other factors that may cause the speakers to need more time. Clark and Wasow (1998) report broad evidence that speakers are more likely to suspend speaking the more complex the following constituent is (see also Kidd et al. 2011 for the relation between complicated, infrequent or discourse- new referents and hesitations).

Stage III in Table 5.1 is where a certain level of speaker agency may be relevant as Clark and Fox Tree (2002) report: speakers are able to anticipate the duration of an expected delay, and accordingly choose different means of dealing with it. In the Qaqet pear stories, speakers deal with the potential delay in speaking by holding level pitch and glottalising the constituent just before the pause, occasionally lengthening the last constituent.[1] This absence of hesitation particles in the pear story data supports a hypothesis formulated by Clark and Fox Tree (2002): that hesitation particles are not a 'symptom of trouble' but rather an interjection that comments on the actual performance of the speaker in announcing 'minor or major expected delays' in formulation (Clark & Fox Tree 2002: 79). As a symptom of processing trouble, it is to be expected that it can be found in all speakers of the world. As a linguistic sign, on the other hand, there may well be typological differences: in the languages described by Streeck (1996), lengthening is used as a hesitation sign, whereas in Qaqet, final level pitch and glottalisation are used (Hellwig 2019: 56).

Stage IV from Table 5.1, the repetition of the preposed function word, restores the continuity of the utterance. It is up until this step that self-interruptions and hesitations are identical. It is true for both that speakers need additional processing time and they handle their delay in some way, but in hesitations, they continue the previous intonation contour, and potentially try to restore its continuity, for example, by repeating the function word. Self-interruptions, conversely, are those disfluency pauses where the intonation contour is abandoned and, accordingly, no continuity has to be restored. Speakers can continue an intonation contour only until some 'upper limit' of time (Himmelmann 2014: 936), so the reset in intonation is rather a consequence of processing issues than of functional differences. This is why, for the following comparison of CDS and ADS in the next section, both phenomena are treated alike.

1 Himmelmann (2014: 942) reports for Tagalog (Philippines) and Streeck (1996: 208) for Ilokano (Philippines) and Lauje (Sulawesi, Indonesia) the lack of hesitation particles. Tagalog, Ilokano and Lauje all belong to the Malayo-Polynesian languages.

## 5.3 Disfluency pauses: Comparison of CDS and ADS

In this section, the two registers CDS and ADS will be compared in terms of disfluency pauses. For each speaker, two values have been calculated: disfluency pauses per 100 words in ADS and disfluency pauses per 100 words in CDS. The results are shown in Figure 5.1.

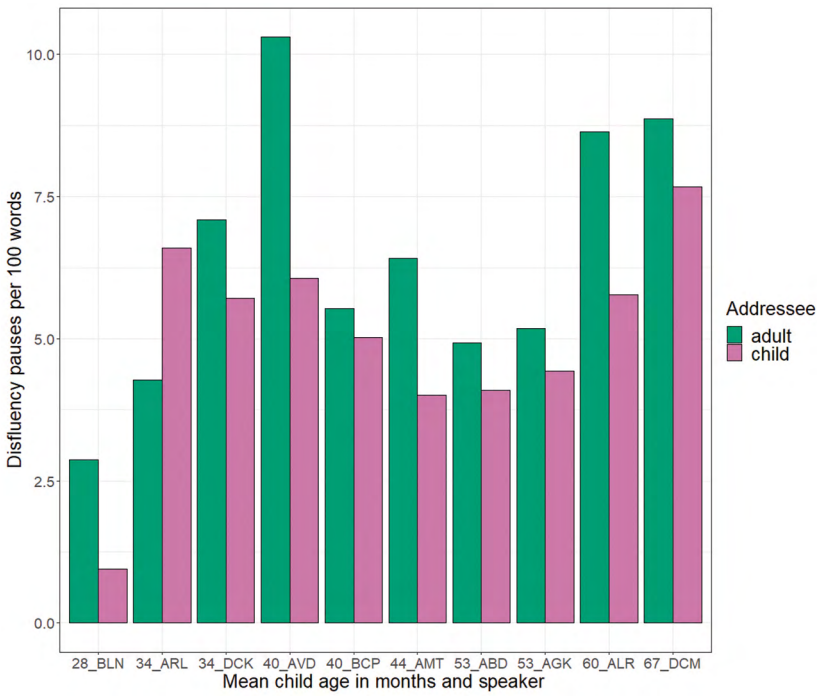


**Figure 5.1: Disfluency pauses per 100 words ADS vs CDS.**

Each pair of bars represents one speaker. The age of the child spoken to in months and its name code are given on the x-axis. On the y-axis, hesitations per 100 words are shown. All but one speaker is more fluent when speaking to a child. Only ARL hesitates more when talking to his son XAT (34m). The mean for ADS is 6.411 disfluency pauses per 100 words (SD = 2.316), and for CDS 5.033 disfluency pauses per 100 words (SD = 1.840). Due to the small size of the sample, a Wilcoxon-test for related samples was chosen. It yielded a significant difference in disfluency pauses per 100 words between ADS and CDS (Z = 48.000; p = 0.037).

This is consistent with previous research that describes a lower disfluency rate in CDS than in ADS. On the basis of these results, I expected that the proportion of disfluencies in CDS correlates with the age of the child, but a non-parametric Spearman test showed that this is not the case ($r = 0.676$, $p = 0.152$). There is therefore no clear connection between child age and CDS fluency. An explanation for these results is offered with reference to individual differences between speakers. These differences are large (see Figure 5.1). Speaker BLN was by far the most fluent speaker in CDS and in ADS, while speaker AVD hesitated a great deal in ADS and lay in the normal range for CDS. Furthermore, the hesitations were occasionally caused by the interactive style of the child. Out of the 28 times speaker DCK hesitated in the CDS-condition, 11 instances occurred when he saw his daughter ZEA look in another direction. At such times, he tried to get her attention by leaning over to her, calling her or tickling her shoulder. Similar evidence is reported by Nilsson Björkenstam et al. (2013) who find that many disfluencies are used as attention-directing signals.

For ARL, a closer examination of the interaction between the only adult speaker (a father) who hesitated more when talking to a child (XAT) helps to explain the results. XAT was the most interactive child in the pear story task, asking more questions than all the other children in the study (see example (39)).

(39) a. ARL: *iakamariaqiamabaketki*
ia-ka=mat
another-3SG.M.SBJ=take/pick_up
ia-ki ama=baket-ki
another-NC.SG.F ART=bucket-NC.SG.F
'The other one takes another bucket…'

b. XAT: *baketkiqua?*
baket-ki kua
bucket-NC.SG.F where/why
'Where is the bucket?'

c. ARL: *deqamuqi daa.. ilanyit braqi*
de=ka=mu-ki de=aa
CONJ=3SG.M.SBJ=put-3SG.F loc.part=3sg.m.poss
ilany-it pet-ki
foot/leg-NC.SG.LONG on/under-NC.SG.F
'and he puts it onto his.. bicycle'

d. XAT: *aailanyit braqiqua?*
aa=ilany-it
3SG.M.POSS=foot/leg-NC.SG.LONG
pet-ki kua
on/under-3SG.F where/why
'Where is his bicycle?'

e. ARL: *de.. iva de.. keuaisaqi*
de ip-a de ke=uaik
CONJ CONJ-DIST CONJ 3SG.M.SBJ.NPST=run
se-ki
to/with-NC.SG.F
'and.. afterwards.. he leaves with it.' (PearARLP 065-070)

Instead of listening to the story, XAT often interrupted his father by asking for the location of things ARL mentioned during the story. ARL had to react constantly to the actions of his son and accordingly lost the thread of his narration. It is likely that this interactive style led to more hesitations, for example, in the case of (39e). The individual style of interaction between caregivers and children again plays a central role.

Still, for all speakers but one, CDS was more fluent than ADS. I next investigated whether the higher complexity of ADS causes its disfluency (see Section 5.1). I compared the MLU of all utterances with hesitations to the MLU of all utterances without hesitations.[2] The results for ADS are presented in Figure 5.2, and for CDS in Figure 5.3.

In ADS, the mean length of utterances with hesitations is 10.93 words (SD = 4.123), while utterances without hesitations have a mean length of 6.35 (SD = 3.620). For CDS, the corresponding values are 10.43 (SD = 3.707) words for utterances with hesitations and 5.18 (SD = 3.617) words for utterances without hesitations. This supports the hypothesis that utterance complexity in terms of utterance length is responsible for the higher rate of hesitations in ADS.

2 Note that self-interruptions are excluded from this count; only full intonation units are part of the calculation.

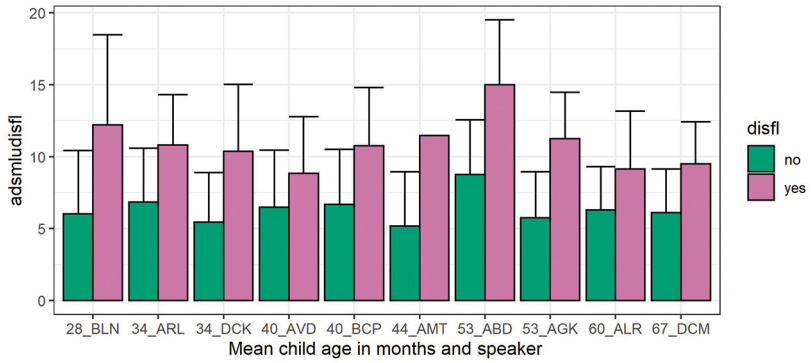


**Figure 5.2: MLU of utterances with and without hesitations in ADS.**

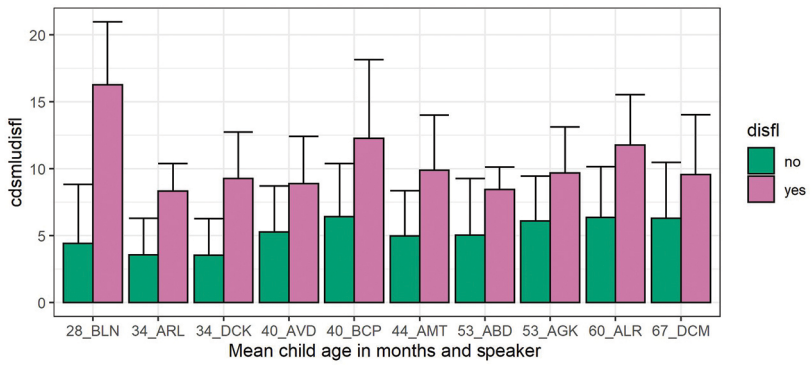


**Figure 5.3: MLU of utterances with and without hesitations in CDS.**

## 5.4 Distribution of hesitations

In Section 1.2 I explained that the location of a hesitation may tell a listener what to expect. To investigate this hypothesis, I annotated the stories (CDS and ADS) of those persons talking to the younger children (of 36 months or below) to determine the position of the hesitation: does it occur before a syntactic phrase boundary, in a situation in which the speaker has to plan what to say next or before a new referent is introduced?[3] For self-interruptions, it is difficult to judge whether the constituent after the pause is what makes people hesitate. As the original

3 Previous studies reporting hesitations at potentially helpful positions (Kidd et al. 2011) have child participants up to three years of age, so I excluded the data from the older children.

utterance is abandoned, the constituent after the pause is not necessarily uttered. Accordingly, only hesitation pauses have been included in the following analysis.

As a matter of fact, in the pear stories, the speakers hesitate often because they are trying to remember what happened in the film. This is typically what happens at the beginning of the stories (see example (40)). ARL has just started to tell the pear story to his friend ACP, but has not yet figured out how to start his story.

(40) *amasiitka de tikai..*

| | | |
|---|---|---|
| ama=siit-ka | de | tika=i |
| ART=story-NC.SG.M | CONJ | EMPH=CONJ |

'the story, so..' (PearARLA 001)

*iak deqatikaqe..*

| | |
|---|---|
| ia-ka | de=ka=tika=ke |
| another-NC.SG.M | CONJ=3SG.M.SBJ=EMPH=3SG.M.SBJ.NPST |

'a man, he..' (PearARLA 002)

Example (40) seems to be a typical instance of what Chafe calls 'finding the focus of consciousness'. He distinguishes this from a second reason to hesitate:

> the speaker's need to find the next focus. Others, as we will see at the end of this discussion, stem from the need to find the best way to verbalize a focus, once found. In other words, sometimes speakers hesitate while they are deciding *what* to talk about next, and sometimes they hesitate while they are deciding *how* to talk about what they have chosen. (Chafe 1985: 171)

For some people from Raunsepna, there are a considerable number of unusual referents in the pear film. There are no goats in Raunsepna, nor pears or bicycles (children have probably never seen one). The goat especially receives many labels (from 'thingy' to 'cow', 'horse', 'dog' and 'donkey') and people hesitate frequently before referring to it (see (41)).

Here, ABD first does not know how to describe the sound the man hears ('He hears the sound of.. well some thingy'), then she explicitly comments that she recalled the word ('They call it goat'). In the following description, she still hesitates before uttering the word *goatkia*. She obviously knows what she wants to talk about, but is searching for the right word to express her thoughts.

(41) a. ABD: *kanarli samerama.. kerl amamaqia*
ka=narli se=met=ama
3SG.M.SBJ=hear/feel to/with=in=ART
kerl ama=ma-ki=a
DEONT ART=thingy-NC.SG.F-DIST
'He hears the sound of.. well of some thingy'

b. ABD: *iratiski amagoatki […]*
i=ta=tis-ki
CONJ=3PL.SBJ=callsay-NC.SG.F
ama=goat-ki
ART=goat-NC.SG.F
'They call it goat […]'

c. ABD: *kiqerl iakekiurlet meraa.. aagoatkia*
kiqerl ia-ka=kiurlet
EMPH another-NC.SG.M=pull
met=aa aa=goat-ki-a
in=3SG.M.POSS 3SG.M.POSS=goat-NC.SG.F-DIST
'A man pulls his.. his goat'(PearABDA 035-038)

In Section 5.1 I reported previous research that found hesitations to be predictive of unknown or difficult referents and phrase boundaries, especially in CDS. Example (41) is one instance of such a situation. Other hesitations precede clausal and phrasal boundaries, see (42). Those might help children identify syntactic structures.

(42) *sagel luqia draawilwilki ip.. ip kirlguirl*
se=gel lu-ki-a
to/with=close DEM-NC.SG.F-DIST
de=araa=wilwil-ki ip
LOC.PART=3PL.POSS=bicycle-NC.SG.F CONJ
ip ki=rlguirl
CONJ 3SG.F.SBJ.NPST=return
'[the man whistles] to the woman on her bicycle so that.. so that she will come back' (PearABDP 091)

Occasionally, a hesitation is influenced by lexical choice (see (43)). BLN hesitates when she wants to refer to the goat and calls it *danggi* 'dog' instead of *goatki* or, the term she uses when talking to her husband, *amamaqi taquarl amahoski* 'Something like a horse'. During transcription of the relevant example, BLN commented that she decided to talk about

an animal she knows her son, ZDL (28), is familiar with. So in this case, the adaptation of the input to the child interlocutor produced a hesitation pause, allowing BLN to select the appropriate referent.

(43) *[…] katden kanaa.. aadanggi*
ka=tden ka=ne=aa
3SG.M.SBJ=come 3SG.M.SBJ=from/with=3SG.M.POSS
aa=dang-ki
3SG.M.POSS=dog-NC.SG.F
'… He comes with his.. his dog' (PearBLNP 031)

## 5.5 Summary: Disfluencies in Qaqet CDS

In this chapter, I compared the number of disfluency pauses (both self-interruptions and hesitation pauses) in ADS and CDS. Nearly all of the hesitation pauses involve a function word uttered before the pause. There are significantly more disfluency pauses in ADS than in CDS, both for fathers' and mothers' speech, but the difference does not correlate with child age. These results confirm previous research proposing a high degree of fluency as typical for CDS. The difference in length between utterances with and without MLU indicates that the lower complexity of CDS is responsible for its fluency.

Additionally, the results show that fluency (at least with reference to hesitations) is not directly related to comprehensibility. Both in CDS and ADS, uncommon referents provoke hesitations. Kidd et al. (2011) found that even small children are sensitive to hesitations announcing complicated referents. There are also hesitations at phrase boundaries, possibly enabling the listeners to locate those boundaries.

Several hesitations in the data are related to the interaction between adult speakers and child listeners. Hesitations are often provoked, for example, because children do not listen and the speakers interrupt themselves to get their attention back, or a child is too interactive and the speaker, busy responding to his or her questions, hesitates as he tries to find his way back to the story. Sometimes, the linguistic adjustments used to adapt speech to children's needs provoke hesitations. Speakers may need additional planning time for the relevant adjustments.

The evidence presented in this chapter suggests that the effects of interaction are more relevant than age for fluency in CDS. This factor could be controlled for by excluding those disfluencies that are caused by interactional features. Still, disfluencies are potentially helpful as cues for referential intentions and identification of salient structures. This has also been proposed for the typical prosodic features of CDS, which are the topic of the next chapter.

This text is taken from *Child-directed Speech in Qaqet: A Language of East New Britain, Papua New Guinea*, by Henrike Frye, published 2022, The Australian National University, Canberra, Australia.

doi.org/10.22459/CDSQ.2022.05