Turn-timing in Norwegian Sign Language

A study of transition durations in question-answer sequences

LING350

Master's thesis in linguistics

Department of Linguistic, Literary and Aesthetic Studies

Astrid Næss Bjørgul

2022

Abstract

Turn-taking requires collaboration between interlocutors, and previous research has found that there is a desire for minimal overlap and gap in informal conversations. Because there is limited research on this topic in signed languages, this thesis investigated turn transition durations in question-answer sequences in informal, Norwegian Sign Language (NTS) conversations. By analyzing a selection of files from two data sets within the Norwegian Sign Language Corpus, the aim was to find out whether mean transition durations in NTS are in the range observed for other spoken and signed languages, if transition durations are variable between individuals, and if age or question type affect mean transition durations. As this study relied on previously collected data the participants were not recruited for this research specifically. The transition durations of 159 question-answer sequences were measured in terms of stroke-to-stroke turn boundaries and yielded 100 gaps and 59 overlaps. The results were in line with previous research on turn-timing, measuring a mean turn transition duration within 250 ms of the cross-linguistic average observed for spoken languages, supporting the theory that turn-timing varies very little across languages, no matter the modality. Some individual differences could be observed, but no significant difference was found between question types, nor between age-ranges, due to few examples in each category.

Sammendrag

Turtaking i uformell samtale krever samarbeid fra samtaledeltagere, og ifølge tidligere forskning er det minimalt med overlappende tale eller lange pauser mellom turer i verbale språk. Det er det imidlertid lite forskning om turtaking i tegnspråk, og denne oppgaven undersøker hvor lang tid turskifter i spørsmål-svar-sekvenser tar i uformelle samtaler i norsk tegnspråk (NTS). Dette ble gjort ved å analysere filer fra to forskjellige datasett i et norsk tegnspråkkorpus med mål om å finne ut om gjennomsnittstider for turskifte i NTS er innenfor den samme tidsrammen som det som er observert for talespråk og andre tegnspråk. Denne oppgaven ser også på om gjennomsnittstider for turskifte varierer mellom individer, og om alder på den som stiller eller svarer på spørsmålet, og spørsmålstype påvirker gjennomsnittstiden. All data brukt i denne oppgaven er hentet fra tidligere innsamlet data, og informantene ble dermed ikke rekruttert spesifikt for dette studiet. Turskiftetiden til 159 spørsmål-svar-sekvenser ble målt i henhold til 'stroke-to-stroke'-turavgrensninger, hvilket medførte 100 pauser og 59 overlapp. Resultatene i denne studien støtter tidligere forskning på gjennomsnittstider i turskifte med en gjennomsnittstid 250 ms innenfor den tverrspråklige tiden på 208 ms observert for ti forskjellige talespråk. Dette støtter teorien om at turskiftetider varierer svært lite i forskjellige språk, til tross for språkets modalitet. Det ble observert noen individuelle forskjeller i gjennomsnittstider, men det var ingen signifikant forskjell på alder eller type spørsmål, trolig grunnet et begrenset datasett.

Acknowledgements

First and foremost, I would like to thank my supervisors, Vadim Kimmelman at The University of Bergen and Lindsay Ferrara at The Norwegian University of Science and Technology, for their guidance through this process.

I want to thank Vadim for being patient with me as a beginner in the field of signed language research, and for helping me with statistics. I also want to thank him for great advice and constructive criticism and for always making me feel better after a meeting.

I want to thank Lindsay for providing me with my data and for helping me through the annotation process. I also want to thank her for the opportunity to present my research to her students at NTNU.

I also want to thank Anita for being a great Norwegian Sign Language teacher and helping me get a better understanding of the language through a fun learning environment.

Lastly, I want to thank Annette and Ingrid for all the laughs, lunches and way too long breaks from writing. I could not have done this without you.

Bergen, May 2022
Astrid Næss Bjørgul

Contents

Abstract .... **II**

Sammendrag .... **III**

Acknowledgements .... **IV**

1 Introduction **1**

1.1 Turns and turn allocation *1*

1.2 Turn-taking research *2*

1.3 Turn-taking in signed languages *3*

1.4 Turn-taking research in Norwegian Sign Language *3*

1.5 Research questions *4*

1.6 Thesis outline *4*

2 Signed languages and NTS **6**

2.1 Signed Languages *6*

2.2 NTS *7*

3 Theoretical Background **9**

3.1 Turn-taking research in spoken languages *9*
3.1.1 The Conversation Analysis model 9
3.1.2 Challenges to the CA model for turn-taking 10
3.1.3 Overlaps and gaps in turn transitions 11
3.1.4 Turn transition durations in spoken languages 13
3.1.5 Summary 14

3.2 Turn-taking research in sign language conversation *14*
3.2.1 Turns, TCUs, and TRPs 15
3.2.2 Overlaps and gaps in turn transitions 16
3.2.3 Transitions durations in signed language conversation 19
3.2.4 Turn-taking in Norwegian Sign Language Research 20
3.2.5 Summary 21

4 Methodology **23**

4.1 Corpus linguistics and sign language corpora *23*

4.2 The Norwegian Sign Language Corpus *24*

4.3 Data and participants *25*

4.4 ELAN *27*

4.5 Movement phases and turn boundaries *29*

4.6 The annotation process *30*
4.6.1 Difficult cases 36

4.7 The data set *40*

4.8 Inter-rater reliability agreement *41*

4.9 Method of analysis *42*

5 Analysis and results **43**

5.1 Quantitative analysis and results *43*
5.1.1 Individual differences 44

5.1.2 The effect of question type and age 48
5.2 Qualitative analysis and results *51*
5.2.1 Gaps 52
5.2.2 Overlaps 60
6 Discussion 65
6.1 Transition durations in NTS 66
6.2 Individual differences 67
6.3 The effect of question type and age 67
6.4 Gaps and overlaps exceeding 1000 ms 68
6.5 Limitations of the study 69
7 Conclusion 71
References 73
Appendix 76

1 Introduction

Conversation is, for most people, an everyday action. We talk back and forth to each other and time our responses with precision, almost effortlessly, and without thinking about how we do it. Though we may not think about it, we use the turn-taking system, a system that changes depending on the context of the conversation. The turn-taking system used in classroom settings is not the same turn-taking system used in formal debates, or in informal conversations. There are sets of rules for each of these systems that distinguish them from one another. In formal debates, for instance, there is someone whose job it is to give out and stop turns. In traditional classroom settings, the teacher manages the turns, and can choose to give students turns. In informal conversation, however, there is no such person whose role it is to manage the turns. Rather, there is a collaboration between each conversation participant to manage the turns between them. This is what makes turn-taking in informal conversation especially interesting to study and this type of turn-taking is what this thesis is about.

Signed languages have been met with an overwhelming amount of prejudice over the years, and even though research dating back to the 1960s has provided evidence of signed languages being natural languages just like spoken languages (Stokoe, 1960), there are still people who believe that signed languages are comprised of mere gestures representing spoken language. A surprising amount of people also believe that there is but one sign language in the world, and that signers all over the world can communicate with each other using one common language. The fact that these myths are still believed suggest that signed languages receive very little attention. For this to change, it is important to investigate all linguistic aspects of signed languages, in as many different signed languages as possible. This includes turn-taking. Very little research has been done on the turn-taking system for informal conversations in signed languages so far, however, and the goal of this thesis is to add knowledge to this topic by looking at turn-timing in Norwegian Sign Language (NTS) using the Norwegian Sign Language Corpus (Ferrara & Ringsø, 2021; Ferrara & Bø, 2015).

1.1 Turns and turn allocation

According to Sacks et al. (1974), the turn-taking system consists mainly of turns and turn allocation. The turn itself is the contribution of a speaker or a signer in a conversation, whereas turn allocation deals with the selection of the next turn. The turn can be broken down into smaller parts, the turn construction unit, and the transition relevance place. The turn construction unit is the part of the turn that ends in a transition relevance place, i.e., the first possible place for a potential new turn, and where turn allocation comes into play. At transition relevance places, the first speaker or signer can either select someone for the next turn, or someone else in the conversation can self-select for that next turn. The first speaker can also self-select and continue beyond the transition relevance place (Sacks et al., 1974, pp. 702–703).

As mentioned above, the timing of responses is very precise. This indicates that the planning of the next turn happens before the transition relevance place. Previous research on turn-timing supports this theory, as the time between turns have been found to average 208 ms across 10 different spoken languages (Stivers et al., 2009). Stivers et al. (2009) states, for reference, that 208 ms is approximately the same time as it takes to produce a single English syllable (Stivers et al., 2009, p. 10588).

1.2 Turn-taking research

Research on turn-taking in informal conversation started in the 1970s when Sacks, Schegloff, and Jefferson described the system for the first time (Sacks et al., 1974). They based their research on recordings of informal speech and described a system in which there is a desire for minimal overlaps and gaps. They also found that "overwhelmingly, one party talks at a time" (Sacks et al., 1974, p. 706). Though it may seem logical, their findings of minimal gap and overlap and that one conversation participant speaks at a time, was the first time this had been investigated in scientific research. This created opportunities for further linguistic research on conversation analysis. It led to research on the timing of turn transitions and the universality of it (de Vos et al., 2015; Stivers et al., 2009), and to research on repair mechanisms during overlapping talk (Schegloff, 2000; Skedsmo, 2020).

Stivers et al. (2009) investigated the timing of turn transitions in question-response sequences in 10 different spoken languages and found a cross linguistic average of 208 ms. More importantly, they found that each individual language had a transition average within 250 ms of the cross-linguistic average, demonstrating that the difference in turn-timing between the 10 languages is miniscule. The average cross-linguistic gap of 208 ms also indicates that there is a desire for minimal overlap and gap, supporting the findings of Sacks et al. (1974). Previous research has also found that response times vary depending on question type, where alternative questions and yes/no-questions get a faster response than wh-questions (Strömbergsson et al., 2013). When it comes to dealing with the overlapping talk that does occur, there is evidence suggesting that there are several repair mechanisms in place (Schegloff, 2000). Schegloff (2000) also found that most of the overlapping talk both happens at transition relevance places (first possible place for a potentially new turn) and that overlapping talk is resolved quickly (for more details, see Chapter 3).

1.3 Turn-taking in signed languages

Research into turn-taking in signed languages started much later than for spoken languages, and even today, it amounts to only a fraction of the total research that has been conducted on turn-taking. However, some studies have been done and can give an indication of whether the turn-taking system in informal conversations works the same way in signed languages as in spoken languages.

Evidence suggests that, just like Schegloff (2000) found in American Sign Language, in Brazilian Sign Language (Libras), there are repair mechanisms in place for preventing or ending overlapping signing as well (McCleary & de Arantes Leite, 2012). There is also evidence that overlapping signing most often occurs at the transition relevance place, and that overlapping signing is very brief (McCleary & de Arantes Leite, 2012). As for the timing of turn transitions, it has been found that when turn boundaries are measured from stroke-to-stroke (see detailed information in Chapter 3), Sign Language of the Netherlands falls within a 250 ms range of the cross-linguistic average of 208 ms found in Stivers et al. (2009) (de Vos et al., 2015). Though these findings are from different languages, and regard different aspects of turn-taking, they give the indication that signed languages may contain all the same characteristics for turn-taking as spoken languages.

1.4 Turn-taking research in Norwegian Sign Language

This thesis investigates turn-timing in Norwegian Sign Language, and linguistic research in NTS is scarce, especially in terms of conversation analysis and turn-taking. However, there is some recent research worth mentioning. Ferrara (2020) looked at the different meanings of pointing in NTS and found that interactional pointing was most often used in turn-taking to manage turns (Ferrara, 2020, p. 10). Skedsmo (2020), on the other hand, investigated other-initiations of repair in NTS using data from an NTS conversation corpus and found that these repairs are produced in more restricted formats than other "preceding repair initiations within the same multiple OIR sequence" (Skedsmo, 2020, p. 560). Though these studies are not directly linked by topic to the current study, all linguistic research on NTS contributes to a broader understanding of the language and how it works.

1.5 Research questions

The background information on turn-taking in informal conversation mentioned above demonstrates that turn-taking may be a system found in all spoken languages. However, there is not nearly enough evidence of this system in signed language research. Though previous studies on turn-taking in signed languages have shown the same trends as in spoken languages, the research methods in the existing literature vary making it challenging to compare results. Another thing to note is that very few signed languages have been investigated in terms of turn-taking in general, and even fewer in terms of turn-timing. To generalize about languages or language types, an abundance of evidence supporting those generalizations must be presented. The purpose of this study, therefore, is to investigate turn-timing in NTS, and to add to existing literature in terms of finding out whether NTS shows the same trends regarding the duration of turn transitions. Specifically, this study will look at the stroke-to-stroke turn boundaries in transition durations of question-answer sequences. The four research questions for this study are thus the following:

1. What is the average duration of transitions, is it in the range observed for other spoken and signed languages?
2. Is average duration of transitions variable between individuals?
3. Is average duration of transitions affected by question type?
4. Is average duration of transitions affected by the age range of the participants?

1.6 Thesis outline

This thesis is divided into 7 chapters. Chapter 1 has outlined background for the research questions and defined key terms. The remainder of the thesis will be organized as follows:

Chapter 2 will give an overview of signed languages and NTS research, Chapter 3 will give an overview of previous research and theories on turn-taking both in spoken and signed languages, Chapter 4 will describe the research methods used in this study, including the data collection, participants, annotation process, and method of analysis. The results and analysis will be presented in Chapter 5, and their implications will be discussed in Chapter 6. Finally, a summary and concluding remarks will be given in Chapter 7.

2 Signed languages and NTS

2.1 Signed Languages

We know today, at least in the linguistic world, that signed languages are natural languages that were not invented by anyone, but rather languages that "develop spontaneously wherever deaf people have an opportunity to congregate and communicate regularly with each other" (Sandler & Lillo-Martin, 2001, pp. 1–2). However, in 1960, a time when language and speech were considered one and the same (Armstrong & Karchmer, 2009, p. 390), it was a big leap forward in linguistic research when William C. Stokoe described American Sign Language as a natural language with all the same characteristics as spoken language (Stokoe, 1960). At this time, it was believed that deaf people would not achieve their full potential through signing and the educational goal for deaf people was to acquire speech (Armstrong & Karchmer, 2009, p. 389). After realizing that signed languages had all the same characteristics as spoken languages, and that it had the same potential for communication (Armstrong & Karchmer, 2009, p. 390), Stokoe formulated a descriptive system of American Sign Language. This helped convince both the academic world and the education system, but also the public that signed languages are natural languages, in which deaf children should be allowed to be taught and communicate (Armstrong & Karchmer, 2009).

Since the 1960s, research on signed languages have, as Sandler and Lillo-Martin (2001) put it, "made a significant contribution to our understanding of human language – its structure; its acquisition by children; its representation in the brain; and its extension beyond communication, in poetry" (Sandler & Lillo-Martin, 2001, p. 1). By studying languages with a different modality than spoken languages, it is possible to demonstrate that there are certain properties all natural languages share, such as grammatical structures for instance and that theories about natural languages should extend beyond speech (Sandler & Lillo-Martin, 2001, p. 1).

Research on signed languages have been done on all sorts of linguistic topics. This is demonstrated in Pfau et al.'s (2012) handbook on signed language linguistics which include topics in phonetics, phonology, prosody, morphology, syntax, semantics, pragmatics, communication in the visual modality, psycholinguistics, neurolinguistics, variation, and change, applied issues, and handling sign language data, are presented. The topics are also addressed from a cross-linguistic (including research on over 40 different signed languages), cross-modal (describing similarities and differences from spoken languages), and theoretic perspective (Pfau et al., 2012). In other words, the linguistics of signed languages are well represented in the book.

Though it is well accepted today that signed languages are natural languages, there are still several common myths regarding this topic, as mentioned in the introduction. One such myth is that signs are gestures, and merely visual representations of the corresponding words in the surrounding spoken languages (Næss, 2021, p. 217). As language contact may occur anywhere there are two or more languages being used, signed languages may be influenced by languages spoken in the same areas. However, as mentioned above, natural languages emerge spontaneously among a community of deaf people and are not invented, and though many also believe that there is only one international sign language (Næss, 2021, p. 217), there is enough research on different signed languages available to prove that, just like in spoken languages, there is great variation (Perniss et al., 2007; Pfau et al., 2012).

Misconceptions may arise when there is limited knowledge about the topic. It is therefore important to continuously work to increase that knowledge, especially when it comes to minority languages, as increased scientific attention will likely increase the status of the language, which again will lead to more interest.

2.2 NTS

NTS has approximately 16,500 signers where around 5,000 are deaf signers (NDF, n.d.). Though the use of NTS has been documented to date back 200 years and the first school that taught through NTS was established in 1825 (Erlenkamp, 2011), the language was only recognized as the national Norwegian sign language in 2021 (Språkloven, 2021, § 7). Before this, the language was not considered a natural language on the same level as spoken languages. About 20 years after Stokoe (1960) described American Sign Language, however, NTS was recognized in linguistic research in the 1980s when it was first described in terms of a "proper" language on the same level as spoken languages by Marit Vogt-Svendsen (Erlenkamp, 2011). At a time with considerable focus on the combination of signed language and speech in teaching (Arnesen et al., 2008, p. 66) she, as a teacher of deaf students, noticed that using NTS signs with spoken Norwegian grammar was ineffective compared to sign language grammar (Vogt-Svendsen, 1981). This observation that the two grammars were quite different from one another, was a starting point for the change of attitudes towards NTS (Erlenkamp, 2011). Vogt-Svendsen published research based on her own experiences, and since then, there has been a somewhat even flow of research on NTS.

Of the early linguistic works, worth mentioning, is Marit Vogt-Svendsen's study of mouth positions and mouth movements to demonstrate that NTS has its own grammatical features which are different from the spoken Norwegian (Vogt-Svendsen, 1981), and her continued work on non-manual components in interrogative structures (Vogt-Svendsen, 1990). In the 2000s, Kari-Anne Selvik published work on temporal expressions in NTS (Selvik, 2006), and Sonja Erlenkamp on gesture verbs (Erlenkamp, 2009). The academic interest in NTS linguistics has spiked in the last decade, with Erlenkamp (2011), Schroder (2011), Raanes (2011), Selvik (2011), Ferrara and Halvorsen (2017), Ferrara and Nilsson (2017), Ferrara and Ringsø (2017), Ferrara (2020), and Skedsmo (2020), all commenting on different linguistic aspects of the language. As recently as 2020, a 150 page book written by Arnfinn Muruvik Vonen, describing the linguistics of NTS was published, with the goal of reaching a broad audience and enlighten all who are interested in the language to learn how NTS works (Vonen, 2020). It describes the similarities and differences between NTS and Norwegian spoken language and includes chapters on the history of NTS, what is included in a language, phonology and phonetics, different types of signs, morphology, syntax, pragmatics, language acquisition, and language contact (Vonen, 2020).

As mentioned above, increased scientific attention of a language will likely increase interest, and though there seems to be an increased interest in NTS linguistics in recent years, and though NTS is fully acknowledged as a natural language today, research on conversation analysis is scarce, and there is not yet to my knowledge any research on turn-timing in NTS. To broaden the understanding of the language, this thesis is an attempt to close that research gap, or at least make it slightly smaller.

3 Theoretical Background

3.1 Turn-taking research in spoken languages

Research on turn-taking in conversation started with Sacks et al.'s (1974) paper on turn taking organization which laid the foundation of Conversation Analysis (CA). The study described a system for turn-taking and presented major findings such as the one-speaker-at-a-time principle, and the discovery that overlaps in conversation are common but brief. A discovery that has led to many studies of the different aspects of overlap in conversation, which will be presented in the next sections.

3.1.1 The Conversation Analysis model

The most common model for turn-taking is Conversation Analysis (CA). Conversation Analysis is an approach used to research language use and social interaction to identify its underlying structures (Sidnell & Stivers, 2013). The approach emphasizes the importance of the data collection as having to be "records of spontaneous, naturally occurring social interactions rather than, for instance, contrived interactions or those that might occur in a laboratory" (Sidnell & Stivers, 2013, p. 2). The field of study is relatively new, and not introduced by linguists, but sociologists Harvey Sacks, Emanuel A. Schegloff, and Gail Jefferson. Their research on turn-taking (Sacks et al., 1974) lay the foundation for Conversation Analysis and is to this day one of the most cited papers in the category of language (Sidnell & Stivers, 2013, p. 3). One of the most important aspects of CA is the belief that social interaction has a strict and detailed orderliness to it, which I will take a closer look at in the next section.

3.1.1.1 Turns, TRPs, and TCUs in the CA model

"A simplest systematics for the organization of turn-taking for conversation" by Sacks et al. (1974) was, as mentioned above, the first attempt at making a model for the organization of turn-taking in conversation. A turn can be described as the contribution of a speaker in a conversation (Girard-Groeber, 2015, p. 2), but a turn can also be broken down into smaller parts. Sacks et al. (1974) used video recordings of "naturally occurring conversations" (Sacks et al., 1974, p. 697) to investigate whether there is a context-free (defined by Sacks et al., 1974 as insensitive to places, times, and identities of parties in the conversation) systematicity to turn-taking. From what they found, they argue that turn-taking can be described in terms of *the turn construction component* and *the turn allocation component* as well as a set of rules.

The turn construction component or unit (TCU) is described as a unit-type with projectability that ends with a possible completion place called *transition-relevance place* (TRP) (Sacks et al., 1974, pp. 702–703), i.e., the place where the next speaker has the opportunity to initiate a turn. Selting (2000) added to this by arguing that "TCUs and turns are the result of the interplay of syntax and prosody in a given semantic, pragmatic, and sequential context" (Selting, 2000, p. 511). She also argued that "TCUs must be conceived of as the smallest interactionally relevant complete linguistic units in their given context", and that these always end in TRPs unless some "linguistic and interactional resources are used in order to project and postpone TRPs to end of larger turns" (Selting, 2000, p. 512).

The *turn allocation component* is, according to Sacks et al. (1974), techniques used to allocate turns, and they are divided into two groups: techniques where the current speaker selects the next speaker, and techniques where the next speaker self-selects for a turn (Sacks et al., 1974). The set of rules apply to the transition-relevance place and describe how the allocation of a turn is coordinated to minimize gaps and overlaps (Sacks et al., 1974, p. 704).

3.1.2 Challenges to the CA model for turn-taking

Several studies have attempted challenging the conversation analysis model for turn-taking by Sacks et al. (1974). Some of these are discussed in Levinson and Torreira (2015). In their research, they discussed, among others, Duncan (1972, 1974) who proposed the idea of turn-taking signals consisting of different cues. They argued that the turns right before a speaker change almost always ended with one such turn ending cue which became the basis of a turn-taking model "where the turn-taking system is entirely under the control of the first speaker" (Levinson & Torreira, 2015, p. 4). This model contradicts Sacks et al. (1974) where speakers self-select at a transition relevance place. Levinson and Torreira (2015) stated that this view of turn-taking is not quite relevant anymore, but that it shed light on the importance of visual cues and "the coincidence of turn transitions with a number of features of turn construction, prosody, gesture, etc." (Levinson & Torreira, 2015, p. 4).

Another such alternate view is one by Heldner and Edlund (2010) who argued that turn-taking does not have all the systematic properties described by Sacks et al. (1974). They found it problematic that Sacks et al. (1974), for instance, claimed that speakers aim at no gap and overlap, and provided findings suggesting that there is in fact an aim for gaps between turns (Heldner & Edlund, 2010). Their study has, however, been criticized for their somewhat unrealistic definition of gaps in turn transitions, and most other research on this topic still agree that speakers aim at minimal gaps (de Vos et al., 2015; Girard-Groeber, 2015; Levinson & Torreira, 2015; Stivers et al., 2009).

Levinson and Torreira (2015) themselves, found Sacks et al.'s (1974) model too limited and added more constraints and observations to turn-taking and argued that a psycholinguistic model should be developed. A model that should include "temporal constraints that turn-taking imposes on language processing" because "conversational interchange is the core form of language use" (Levinson & Torreira, 2015, p. 13).

3.1.3 Overlaps and gaps in turn transitions

Through the model for turn-taking mentioned above, Sacks et al. (1974) found that "overwhelmingly, one party talks at a time" (Sacks et al., 1974, p. 706), and that some overlap does occur, but it is brief and commonly occurs when two speakers self-select at the same TRP. They also found that transitions with no gap and no overlap are common and that there are repair mechanisms for dealing with errors and violations, such as overlap in turn-taking (Sacks et al., 1974).

Schegloff (2000) argued that the desire for minimal overlap and gaps in conversation is not due to politeness, but due to the fact that "the absence of such an organization would subvert the possibility of stable trajectories of action and responsive action through which goal-oriented projects can be launched and pursued through talk in interaction" (Schegloff, 2000, p. 1). He argued that overlap occurs mostly when two speakers compete for a next turn in "terminal overlap" (Schegloff, 2000, p. 7), and that there are several overlap resolution devices that speakers use to minimize overlap and gaps. For instance, when there is competition to maintain the floor, it is often "negotiated on a syllable-by-syllable basis, with e.g., deceleration, increase of intensity, and repeated syllables or words, until one speaker drops out" (Levinson & Torreira, 2015, p. 3). Schegloff (2000) also states that "the vast majority of overlaps are resolved to a single speaker by the third beat" (Schegloff, 2000, p. 24), demonstrating how quickly overlap is resolved.

When gaps happen in turn transitions, Levinson and Torreira (2015) argued that there are several different categories of "absence of speech" such as gaps between speakers, silences between questions and their responses, and lapses where no speaker self-selects. They stated that long lapses suggest that the speaker responding may find the question problematic in some way and difficult to answer. This may cause the speaker who asked the question to predict the answer, which again may lead to overlap, whether they predict correctly or not. However, they mentioned that these overlaps are rare, but slightly more common if it is clear that the second speaker is searching for a word (Levinson & Torreira, 2015, p. 3).

Levinson and Torreira (2015) analyzed a corpus of 348 dyadic conversations. They used a classification scheme by Heldner and Edlund (2010) which distinguished between two types of overlap: Within-overlap, where speaker 2 starts a turn during the turn of speaker 1 and ends the turn before speaker 1 has ended their turn, and Between-overlap, where one speaker starts a turn during another speaker's turn, and the first speaker drops out (see Figure 3.1).

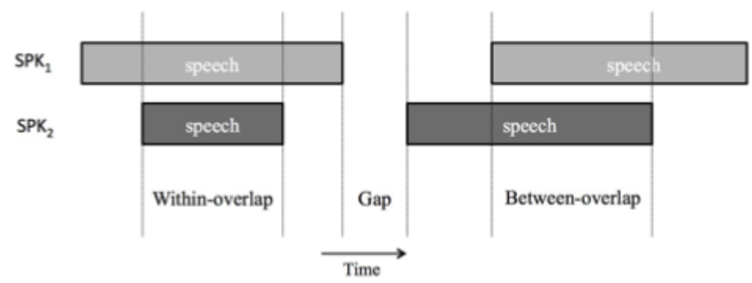


Figure 3.1: Illustration of the different types of overlap discussed in Levinson and Torreira (2015) (Levinson and Torreira, 2015, p. 7).

Their findings showed that speech by one speaker only amounted to 77 percent of the signal, while 19.2 percent corresponded to pauses within a speaker's turn or gaps, and 3.8 percent corresponded to the two types of overlaps (Levinson & Torreira, 2015). In transitions, overlaps corresponded to 30 percent, but they were brief (less than 5 percent of the speech signal). They also found that the overlaps occur mostly in places such as after possible completions or in simultaneous turn-starts, etc. (Levinson & Torreira, 2015). They concluded that the "vast majority of instances of overlap in our dyadic conversations are consistent with the turn-taking system proposed by Sacks et al. (1974)" (Levinson & Torreira, 2015, p. 8).

3.1.4 Turn transition durations in spoken languages

One of the most cited studies on turn transition durations is by Stivers et al. (2009). They undertook a systematic cross-linguistic comparison, where they investigated differences in turn transition durations, using a sample of 10 different spoken languages from five different continents. The languages in question varied in word order, sound structure, grammatical options, etc., (Stivers et al., 2009). They compared data from informal conversations and measured the turn transitions between polar questions and their responses. The findings of Stivers et al. (2009) suggested: "a strong universal basis for turn-taking behavior, in that all languages show a similar distribution of response offsets" (Stivers et al., 2009, p. 10589). The overall mean of the dataset was 208 ms, while the mode was 0 ms, and the median was 100 ms. All languages had an average gap within a range of 250 ms from the cross-language mean (Stivers et al., 2009, p. 10588). This demonstrated that there was a tendency to try to minimize both gaps and overlaps in all the languages included in the study (Stivers et al., 2009). The study only included spoken languages, however, but as we can see later, in subsection 3.2.3, research on Sign Language of the Netherlands suggested the same tendency of minimizing both gaps and overlaps (de Vos et al., 2015) which may indicate that turn-timing in conversation works the same way across language modalities.

Stivers et al. (2009) only included yes/no-questions in their study. After examining a Dutch conversation corpus and investigating timing across all types of turns and responses, they did not find a difference between response times in questions and nonquestions (Stivers et al., 2009, p. 10588). Stivers et al. (2009) might not have found a difference between questions and nonquestions, but Strömbergsson et al. (2013) found that there was a difference in response time depending on the type of question. They investigated Swedish face to face conversations and English phone conversations and found that wh-questions and open questions had a significantly longer response time than alternative questions and yes/no questions. They argue that this might be because in alternative and yes/no questions, response alternatives are provided (Strömbergsson et al., 2013, p. 2587).

Though it seems that most research on turn-timing agrees with Sacks et al.'s (1974) findings that speakers aim at minimal gaps and overlaps, Heldner and Edlund (2010) found it problematic, as mentioned in subsection 3.1.2. In their study, they found that no gaps, which they define as pauses up to 10 ms, are extremely rare, and only accounted for 1 percent of turn transitions in their corpora, while overlapping transitions, on the other hand, accounted for 40 percent (Heldner & Edlund, 2010, p. 562). Heldner and Edlund (2010) argued that the lack of minimal gap and overlap provides evidence that there is no aim to avoid overlap or gaps (Heldner & Edlund, 2010, p. 564). However, Levinson and Torreira (2015) argue against this claim because most conversation analysts define gaps as pauses that last for 150 ms and longer and a pause that is shorter than 150 ms is perceptually not a gap (Levinson & Torreira, 2015, p. 4), i.e., Heldner and Edlund (2010) did not find evidence that does not support the theory of minimal overlaps and gaps. For perspective, they added that voiceless stops in English can last for 60-80 ms (Levinson and Torreira, 2015, p. 4).

3.1.5 Summary

Though there is some disagreement, most researchers maintain that there is an aim to avoid overlap and gaps in conversation and that generally, one speaker contributes at a time. This has been demonstrated by Levinson and Torreira (2015) who found that overlapping talk only accounted for 3.8 percent of the speech signal in their dataset. Adding to this, by measuring the time between yes/no questions and responses, Stivers (2009) found that the aim to minimize overlaps and gaps is consistent across several different languages. When overlapping talk does occur, Schegloff (2000) found that speakers use overlap resolution devices to avoid further overlap (i.e., increase in the intensity of speaking, repetition of words or syllables, etc.). The research presented above shows that there is a systematicity to turn-taking in conversations in spoken languages. In the next sections, we will see that the same systematicity seems to apply to signed languages.

3.2 Turn-taking research in sign language conversation

Research on turn-taking and overlapping talk in signed languages is scarce, and in what exists, the focus is varied. Coates and Sutton-Spence (2001) focused on whether the one speaker at a time principle by Sacks et al. (1974) exists in British Sign Language (BSL), Girard-Groeber (2015) focused on different types of overlap in Swiss-German Sign Language (DSGS) conversation, McCleary & de Arantes Leite (2012) focused on in what environments overlap occur in Brazilian Sign Language, and de Vos et al. (2015) focused on the timing of turn transitions in Sign Language of the Netherlands (NGT) in comparison to spoken language. In addition to varying focus, the research mentioned above also varied in their methodology, and not all researchers agree on where the turn boundaries are. This will all be discussed in this section.

3.2.1 Turns, TCUs, and TRPs

To talk about turns, TCUs, and TRPs in signed languages, we must first touch upon the lexical unit. In signed languages, the lexical unit is the sign, and the sign consists of up to four different movement phases (see Figure 3.2). The phases are preparation, stroke/expressive phase, hold, and retraction. During the preparation phase, the hands move from the resting position into the position of the selected sign. During the stroke, the lexical/meaningful unit is realized, "the form of the body movement is associated with the information to be conveyed" (Kita et al., 1998, p. 28). The hold is the optional holding of the end of the stroke, to for example prompt a response from another conversation participant (de Vos et al., 2015, p. 4). During the retraction phase, the hands move back to the resting position. It is important to highlight that for each sign, only the stroke phase is obligatory, whereas the preparation, hold and retraction phases do not always occur (Kita et al., 1998, p. 27).

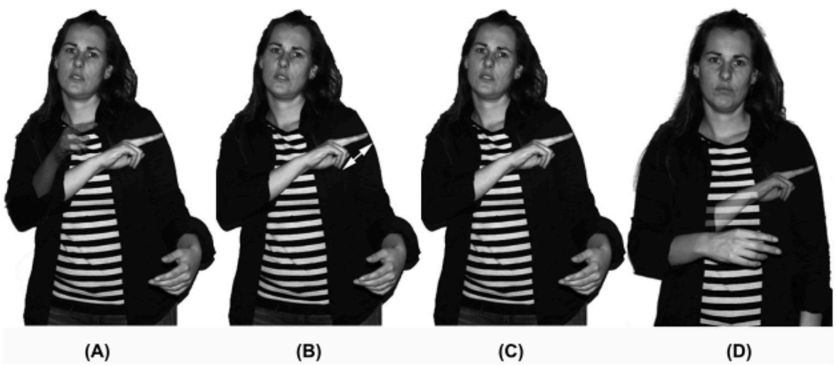


FIGURE 1 | The four gestural movement phases of the Nederlandse Gebarentaal (NGT) sign BROER *'brother'*: (A) preparation, (B) stroke, (C) hold, and (D) retraction.

Figure 3.2: The four movement phases of a sign (de Vos et al., 2015, p. 4).

Because of the complex temporal structure of signs, researchers do not all agree on where the turn boundaries are. Some researchers (Lackner, 2009; Baker, 1977) have stated that the hold of the last sign in a turn is part of the turn and that it is the lifting and lowering of the arms that make up the turn boundaries (Girard-Groeber, 2015, p. 3). However, more recent research by de Vos et al. (2015) adopted the phase coding scheme by Kita et al. (1998), and they argued that "in optimizing turn transitions, signers focus on the phonological content of signs as represented by the stroke, and disregard early preparatory movements, and the intentional holding of signs for response, as well as post-utterance retraction" (de Vos et al., 2015, p. 2). In their analysis, therefore, they included turn boundaries measured from the end of the final stroke in the first speaker's last sign (the TRP), to the beginning of the first stroke in the second speaker's first sign (de Vos et al., 2015, p. 4). McCleary & de Arantes Leite (2012) agreed with this view and state that "we conclude that inbreaths, a recognized practice in preparing the vocal tract for speaking, is canonically understood to constitute preparation for speaking, as opposed to part of the speaking, and thus be legitimately excluded from overlap" (McCleary & de Arantes Leite, 2012, p. 133). They went on to say that "we may, applying the same logic, consider preparation for signing in signed languages, also to be legitimately excludable from overlap" (McCleary & de Arantes Leite, 2012, p. 133).

Girard-Groeber (2015), on the other hand, stated that they include the preparation phase of the initial sign and argued that this corresponds to inbreaths in spoken language which can initiate a turn, though not initiate a TCU yet (Girard-Groeber, 2015, p. 4). However, in their analysis of overlap in signed interaction, they excluded the preparation phase, the hold, and the retraction phase. They argued that this is due to the differentiation between the participants' turn as the overall contribution, and the grammatical units which can occur within that turn, such as a TCU (Girard-Groeber, 2015, p. 4).

3.2.2 Overlaps and gaps in turn transitions

Coates and Sutton-Spence (2001) investigated overlapping turns in British Sign Language. Their research questions were based on a theory by Edelsky (1981) that stated there are two modes of conversation organization: the collaborative floor, where the speakers share the conversation floor, characterized by overlapping talk, and the 'no gap, no overlap' model Sacks et al. (1974) proposed where speakers follow the one-at-a-time-principle (Coates & Sutton-Spence, 2001, p. 507). In their study, they wanted to find out if these two modes of conversation organization apply to signed languages as well, and if so, if there are gender differences. Coates and Sutton-Spence (2001) collected data from informal conversations between friends. The study included two groups of BSL users in their twenties. One group was of females, and one of males. They claimed that signers do not generally follow the one-at-a-time-principle that Sacks et al. (1974) proposed for turn-taking in spoken languages. They concluded that "signers can and do take advantage of a collaborative floor" (Coates & Sutton-Spence, 2001, p. 525), and that female signers more often than male signers choose the collaborative floor (Coates & Sutton-Spence, 2001, p. 526). They even stated that the reason for this is that the risk of their contribution potentially not being seen is outweighed by "the capacity of polyphonic talk to symbolise solidarity and connection" (Coates & Sutton-Spence, 2001, p. 527).

Although Coates and Sutton-Spence (2001) found that the collaborative floor mode of conversation organization is common (i.e., overlapping talk), there are no specific measures or any mention of how they defined the turn boundaries, making the study difficult to compare to other studies on overlap in conversation.

Girard-Groeber (2015) studied different types of overlap in Swiss-German Sign Language (DSGS) based on data from a corpus of DSGS interactions from a previous larger project. The data used in this study is a 33-minute recording of a four-party naturally occurring interaction. It contained 331 overlaps and intended to describe overlaps in terms of what action is accomplished with the overlapping turn, and in what sequential environment the overlapping turn occurred (Girard-Groeber, 2015). They look at three different categories of environments in which overlapping turns occurred. (1) When the overlap occurred at a first possible completion place. (2) When the overlap occurred near the end of a unit, but not quite at the transition place. (3) When the overlap occurred in the middle of a unit (Girard-Groeber, 2015, p. 8). They provided a table for further explanation (see Table 3.1).

Table 3.1: The categories of overlap environments from Girard-Groeber (2015, p. 8).

	Category	Overlap onset with respect to the current signer's turn	Explanation (Participant A = current signer; participant B = incipient/overlapping signer)	Quantification
A	Start of new unit	At a first TRP, i.e., after the stroke of the potentially last sign	Both participant A and participant B launch a unit which is syntactically independent from what precedes (new syntactic construction); this can occur after a pause or straightforwardly after a TRP	110/331 (33.2%)
B	Extension of unit	At a first TRP, i.e., after the stroke of the potentially last sign	Participant B launches a turn and participant A adds one or several signs after his initial TRP; these signs are dependent from the first part of the turn; this can occur after a pause or straightforwardly after a TRP	
C	End of unit	During the stroke of the last item of a turn	Participant A produces the last item of his turn (and then retracts his hands), participant B's turn-initial sign (stroke) overlaps the deployment phase (stroke) of that last item	7/331 (2.1%)
D	Potential end of turn followed by continuation	During the stroke of an item that *could* be the last item of the turn	Participant A produces an item that could be the last item of the turn, participant B's turn-initial sign (stroke) overlaps the deployment phase (stroke) of that potentially last item—but after that item A continues with additional signs or with a new (syntactically independent) unit	146/331 (44.1%)
E	Midst of unit	After the beginning of a syntactic unit, not in potential transition space	Participant B launches a turn while participant A's turn is not yet reaching a possible completion	35/331 (10.6%)
F	Undetermined			33 (10%)
			Total overlaps	331

As mentioned in subsection 3.2.1, Girard-Groeber (2015) excluded the preparation phase, the hold, and the retraction phase from their analysis of overlap.

The major findings of Girard-Groeber (2015) were that overlap occurred most frequently (79.4% of all overlaps) in the environment of possible completion and that they rarely (10.6% of all overlaps) occur in the middle of a unit (Girard-Groeber, 2015, p. 16). They stated that "the majority of overlaps result from the fact that incipient signers anticipate a turn-end and overlap it, and/or the current signers continue beyond a first possible completion" (Girard-Groeber, 2015, p. 16).

Another thing to note is that, where Coates and Sutton-Spence (2001) stated that signed languages are not necessarily driven by the one-at-a-time principle, but rather prolonged simultaneous signing using the 'collaborative floor', Girard-Groeber (2015) argued that the two are not mutually exclusive. They argued that in certain overlaps, speakers still orient to the principle by "keeping their turns short (as in acknowledgments and agreements), or by accomplishing actions that can reasonably overlap ongoing turns because of their urgency (repair initiations) or because of the interactional effects it thereby creates (e.g., strong disagreement)" (Girard-Groeber, 2015, p. 17). However, they did emphasize that their study did not disprove Coates and Sutton-Spence's (2001) proposal that simultaneous 'talk' occurs more often in signed languages than in spoken languages (Girard-Groeber, 2015, p. 16).

McCleary & de Arantes Leite (2012) agreed with Girard-Groeber (2015) that turn-taking in signed interaction is well organized. In their study, they looked closely at turn management in Brazilian Sign Language (Libras). They used excerpts from semi-spontaneous conversations between two fluent signers and analyzed turn-taking in terms of interaction skills in contexts of overlap. They used Schegloff's (2000) study of overlap and overlap resolution devices (ORDs) in spoken conversation as a basis of their analysis and compared and discussed several of these ORDs, and how they work in Libras. They found for instance that signer 1 slowed down the signing pace in response to signer 2's body behavioral display of a "potential intent to initiate a turn" (McCleary & de Arantes Leite, 2012, p. 136), and signer 1's abrupt cut-off and total retraction in response to signer 2's gesture deployment of fixing their hair because this might also be a sort of preparation for an initiation of a turn (McCleary & de Arantes Leite, 2012, p. 136). These are just two examples of many that McCleary & de Arantes Leite (2012) found which indicate that overlap resolution devices in turn-taking work the same way in signed conversation as they do in spoken conversation and that signers also do coordinate their turns with precision.

3.2.3 Transitions durations in signed language conversation

Stivers et al. (2009) found, as discussed in subsection 3.1.4, that in turn transitions in 10 different spoken languages, all gaps were within 250 ms of the cross-language mean (208 ms), and that there was a tendency to minimize both gaps and overlaps. de Vos et al. (2015) wanted to find out to what extent this applies to signed languages as well, specifically to NGT. Following the lead of McCleary & de Arantes Leite (2012) with regards to stroke phases, they chose to exclude the preparation phase, the hold, and the retraction phase from one of their analyses (de Vos et al., 2015, p. 4). They hypothesized that by calculating stroke-to-stoke boundaries, signed conversations should show the same tendencies as Stivers et al. (2009) found for spoken languages (de Vos et al., 2015, p. 2). Their analyses consisted of 190 question-answer sequences from the NGT Interactive corpus which contained recordings of informal conversations by NGT signers. They focused on manual signs, and questions formed with solely non-manual movements were excluded from the analyses (de Vos et al., 2015, p. 3).

de Vos et al. (2015) looked at three different phonetic measures: sign-naïve turn boundaries which included all movement phases of the manual sign, stroke-to-stroke turn boundaries which only included the strokes, and a measure that included the preparation phase for the first sign of the respondent, but only the final stroke for the signer asking the questions (de Vos et al., 2015, p. 4).

They found that for the first measure of sign-naïve turn boundaries, the average offset of the answer to a question was –812 ms, the median was –607 ms, and the mode was –361 ms. For the second measure (stroke-to-stroke turn boundaries), they found an average offset of 307 ms, a median of 269 ms, and a mode of 227 ms. For the third and final measure they found an average offset of –86 ms, a median of –78 ms, and a mode of –53 ms (de Vos et al., 2015, p. 6). The results clearly showed that only one of the three measures, namely the stroke-to-stroke turn boundaries measure, fell within the 250 ms range of the cross-language average found in Stivers et al. (2009). This supports the theory that the preparation phase of the sign closely resembles the pre-utterance inbreath in spoken languages, as McCleary & de Arantes Leite (2012) argued.

The results also indicated that NGT shows the same tendencies for minimizing gaps and overlaps as found in many spoken languages, contradicting Coates and Sutton-Spence's (2001) findings that sign language users do not follow the one-at-a-time principle, but rather use a collaborative floor recognized for overlapping talk. Comparing the two studies, however, is difficult since Coates and Sutton-Spence (2001) did not provide any numerical values or information regarding what they considered to be the turn boundaries which has also been criticized by McCleary & de Arantes Leite (2012).

3.2.4 Turn-taking in Norwegian Sign Language Research

Turn-timing and the study of overlap in NTS has yet to be explored, but some research has been done on certain aspects of turn-taking. Skedsmo (2020) is a qualitative and quantitative corpus study of other-initiations of repair (OIR) in informal NTS conversations. OIR is, according to Skedsmo (2020), self-repair after another speaker in the same conversation has announced a need for it (p. 534) The study presented three different trajectories of multiple OIRs, when one trouble source is targeted by more than one repair initiation, when the self-repair becomes a new trouble source and is targeted by another repair initiation, and when the repair initiation becomes a new trouble source (Skedsmo, 2020). All three trajectories were found in NTS, but the first two were the most common, which was in line with previous research from other languages.

Ferrara’s (2020) study on pointing in NTS also deals with turn-taking. The study examined interactional functions of finger pointing by analyzing video recordings of informal signed NTS conversations (Ferrara, 2020). Of the interactional functions of pointing, it was found that the one most frequently used was the turn-regulating function which amounted to 45.3 percent (p. 16). The turn-regulating functions included giving turns to other signers, often in question-contexts, taking turns, indicate that the turn is free to take, to pause their turn, and to guide the gaze of other signers (Ferrara, 2020, p. 16).

Though this research is not directly relevant to this thesis, in terms of gaps and overlaps, it does demonstrate just how limited the knowledge about conversational turn-taking is in NTS. This thesis intends to add to this existing knowledge and to shed some light on turn transition durations and overlap to create a slightly better understanding of NTS.

3.2.5 Summary

Though research on turn-taking in signed languages still has a long way to go, the research presented above gives a good indication that turn-taking in informal conversation follows the same principles and rules regardless of modality. Girard-Groeber (2015) found that overlapping talk occurred most often at possible completion places, not in the middle of units, supporting the one-at-a-time principle, and McCleary & de Arantes Leite (2012) described overlap resolution devices being used in their dataset, demonstrating that, when overlap does occur, there are ways terminate it. As a comparison to Stivers et al.’s (2009) study on turn transition durations in ten different spoken languages, de Vos et al. (2015) found that, when measuring stroke-to-stroke turn boundaries, the average response offset was 307 ms, i.e., within 250 ms of the cross-linguistic average of 208 ms. Skedsmo (2020) investigated the presence of multiple other-initiations of repair in NTS and presented findings which supported previous studies on the topic in other languages, whereas Ferrara (2020) investigated pointing in NTS and found that turn-taking was the most frequent interaction function.

Even though research on turn-taking in signed languages thus far has yielded results indicating a similar systematicity as has been found for spoken languages, these results are based on very few studies and very few languages and much more evidence from signed languages across the world is needed to conclude anything. This thesis will contribute by adding knowledge about NTS to the topic of turn-timing in signed languages.

4 Methodology

This study examined turn transition durations in NTS, by drawing on data from the Norwegian Sign Language Corpus. The corpus consists of ELAN annotation files with video recordings of informal conversations in NTS. For this study, a selection of these files was used, and stroke-to-stroke turn boundaries were annotated to measure the duration of the transitions between questions and responses. This chapter includes detailed explanations of the methodological process of the study, and the subsections of the chapter are Corpus linguistics and signed language corpora, The Norwegian Sign Language Corpus, Data and participants, ELAN, Movement phases and turn boundaries, The annotation process, The data set, Inter-rater reliability agreement, and Method of analysis.

4.1 Corpus linguistics and sign language corpora

Corpus linguistics is linguistic research based on a body or collection of language. A corpus consists, for the most part, of bodies of texts, but can also consist of transcribed speech or signed conversation (Johannessen & Erlenkamp, 2003, p. 141). The corpora are often tagged for grammatical categories, word classes, etc., depending on what type of texts/speech/signing it consists of and what purpose the corpus has (Johannessen & Erlenkamp, 2003, p. 142). Language corpora can be used for descriptive or theory-based research in grammar, language variation, discourse analysis, etc., and can both be used in synchronic and diachronic studies (Johansson & Oksefjell, 1998, p. 3).

A characteristic of a linguistic corpus is that it consists of a large collection of naturally occurring data, stored in electronic form (Conrad, 2002, p. 76). For a text corpus, the content is naturally occurring by default as it is usually a collection of texts already written before being collected for the corpus. This can be a newspaper corpus or a corpus of Shakespeare's literary work for instance. A speech, or sign language corpus, on the other hand, faces challenges regarding the recording procedure. To gather data for a corpus, the participants must be informed of the research, as well as consent to being recorded, and to the recorded data being used. Though participants may be asked to ignore recording equipment, it may nevertheless influence the data collection to some extent (Johannessen & Erlenkamp, 2003, p. 143).

Another characteristic of a language corpus is that it is searchable to some extent (Johannessen & Erlenkamp, 2003, p. 141). In text corpora, concordance, for instance, can show words in contexts as well as calculate word frequency, analyze words that occur together (collocates), and "often calculate statistical measures of the strength of word associations" (Conrad, 2002, p. 77). In addition to this, Johnston (2009) states that "a modern linguistic corpus contains linguistic annotations and appended sociolinguistic and sessional data (metadata) that describe the participants and the circumstances under which the data was collected" (Johnston, 2009, p. 87). This applies to signed language corpora as well.

Speech corpora and signed language corpora are much less common than text corpora. This is due to challenges such as consent from participants to be recorded, the expense of recording equipment, and the fact that these recordings must be transcribed and annotated manually in order for the corpus to be searchable, which is very time-consuming (Johannessen & Erlenkamp, 2003).

To study informal interaction such as turn-taking in signed conversation, the data should consist of naturally occurring, spontaneous conversation, in informal settings between signers. A sign language corpus is thus well-suited for gathering such language data. At this point, there are a limited amount of signed language corpora available, which is not surprising due to the time-consuming process of creating one, combined with the fact that sign linguistic research is still a very young field. However, in chapter 3, a few corpora, such as The NGT Interactive Corpus and The Corpus of Interactions of Deaf and Hard-of-Hearing Signers in Swiss-German Sign Language, were mentioned, and there are currently researchers working on a Norwegian Sign Language corpus (Ferrara and Ringsø, 2021; Ferrara and Bø, 2015) at The Norwegian University of Science and Technology, which is the corpus used in this study.

4.2 The Norwegian Sign Language Corpus

The data used in this study is taken from the Norwegian Sign Language Corpus, curated by researchers at the Norwegian University of Science and Technology (NTNU). The videos in the NTS corpus are from several data sets, with the earliest being collected in 2012 and the data collection is still ongoing (Ferrara and Ringsø, 2021; Ferrara and Bø, 2015). The annotation process of these videos is in progress, as the corpus is still being developed for further research.

The recordings in each data set were made in various university locations or at local deaf association locations. They were made in informal settings and consisted of spontaneous and "naturally occurring" signing. The conversations are about everyday topics between fluent signers, both hearing and deaf. All data has been ethically approved by the Norwegian Centre for Research Data, and all participants have consented to the data being used in the research (Annotating Norwegian Sign Language corpora, 2022).

Permission was granted by the NTNU researchers who manage and develop the corpus, for this thesis project, to access it and copy the videos and annotation files and further annotate said files independently from the NTS corpus.

4.3 Data and participants

To address the research questions and investigate turn transition durations in question-answer sequences in NTS, a selection of conversations from the NTS corpus were chosen to analyze. In this study, due to limited time and recourses, nine conversations from two data sets from the NTS Corpus were used to collect data: the DPNTS data set (Ferrara and Ringsø, 2021), and The Pilot corpus (Ferrara and Bø, 2015).

The DPNTS data set included 13 conversations with signing from 22 people. The full data set totaled 07:27:40 (seven hours, 27 minutes 40 seconds). Turn-taking during question contexts were investigated in eight conversations from this data set, which included signing from 16 people, nine female signers and seven male signers. Their ages ranged varied from 18 to 60 years. All 16 were deaf signers, fluent in NTS, who had acquired NTS between the ages of 0 and 12 (see Table 4.1).

The Pilot corpus data set included 3 group conversations with signing from 9 people, which totaled 02:01:37. One conversation from this data set was used to investigate turn-taking in question contexts which included three signers and totaled 00:38:41. The conversation included two male signers and one female signer. They were all deaf signers, fluent in NTS, within the age range of 60-79 years, and had acquired NTS between ages 0 and 12 (see Table 4.1).

Table 4.1: Participant background information.

Participant code	Gender	Age ranges	Hearing/Deaf	Age of NTS acquirement
BHS	Female	18-29	Deaf	0-7
TR2	Female	30-40	Hearing	0-7
CJV	Female	18-29	Deaf	0-7
TJ	Male	18-29	Deaf	0-7
EL	Male	30-40	Deaf	0-7
MF	Male	18-29	Deaf	0-7
PS	Female	30-40	Deaf	0-7
EMN	Female	50-60	Deaf	0-7
LPL	Male	18-29	Deaf	0-7
ER	Female	40-50	Deaf	8-12
KFV	Female	50-60	Deaf	0-7
ØSR	Male	30-40	Deaf	0-7
MS	Male	40-50	Deaf	0-7
OIS	Male	60-70	Deaf	0-7
EB	Female	70-80	Deaf	8-12
TR	Male	70-80	Deaf	8-12
PN	Male	18-29	Deaf	0-7
TH	Female	18-29	Deaf	0-7
IMH	Female	18-29	Deaf	0-7

As this corpus relies on previously collected data, the participants were not recruited for this study specifically. However, all participants consented to having the video recordings used in further research.

The nine conversations selected from the two data sets were chosen because, in these, question-related tiers had already been annotated by two NTNU researchers for another study on questions in NTS. The nine conversations used in this study varied in length from 26 to 45 minutes. The number of signers per group also varied, where six of the groups were triads, and three of the groups were dyads. All groups had questions, as mentioned, annotated. However, the number of questions in each conversation varied greatly, from 3 to 62 questions (see Table 4.2).

Table 4.2: Conversation group information.

Conversation number	**Participants**	**Data set**	**Conversation type**	**Recording length (m:s.ms)**	**Number of questions**
1	BHS, CJV	DPNTS	dyad	26:47.739	3
2	CJV, TJ	DPNTS	dyad	31:55.179	52
3	EL, MF, CJV	DPNTS	triad	26:58.083	9
4	EMN, TR2	DPNTS	dyad	40:57.560	8
5	ER, PS, TR2	DPNTS	triad	45:23.080	3
6	KFV, ØSR, MS	DPNTS	triad	28:38.383	62
7	OIS, EB, TR	Pilot	triad	38:41.660	46
8	PN, TR2, LPL	DPNTS	triad	43:27.972	6
9	TH, IMH, KFV	DPNTS	triad	28:31.209	5

Question-answer sequences were chosen for this study because, as de Vos et al. (2015) puts it, "question-answer sequences provide a particularly well-suited conversational context in which to investigate turn-timing, as questions make due a conditionally relevant and timely response" (de Vos et al., 2015, p. 2). It also makes the study comparable to Stivers et al.'s (2009) study on turn-timing in ten different spoken languages and can indicate whether NTS falls within the same cross-linguistic transition-time range (208 ms) as was found in that study.

4.4 ELAN

In this study, ELAN version 6.2 was used (Wittenburg et al., 2006). ELAN is a linguistic annotation platform where text annotations can be created for video and audio files (Wittenburg et al., 2006). The platform allows users to document and analyze signs or gestures by creating *tiers*. A tier is a set of annotations that share the same characteristics (Brugman & Russel, 2009) i.e., an annotation category. When annotations are added, they may be added to a timeline on the specific tier to which it belongs (see Figure 4.1).

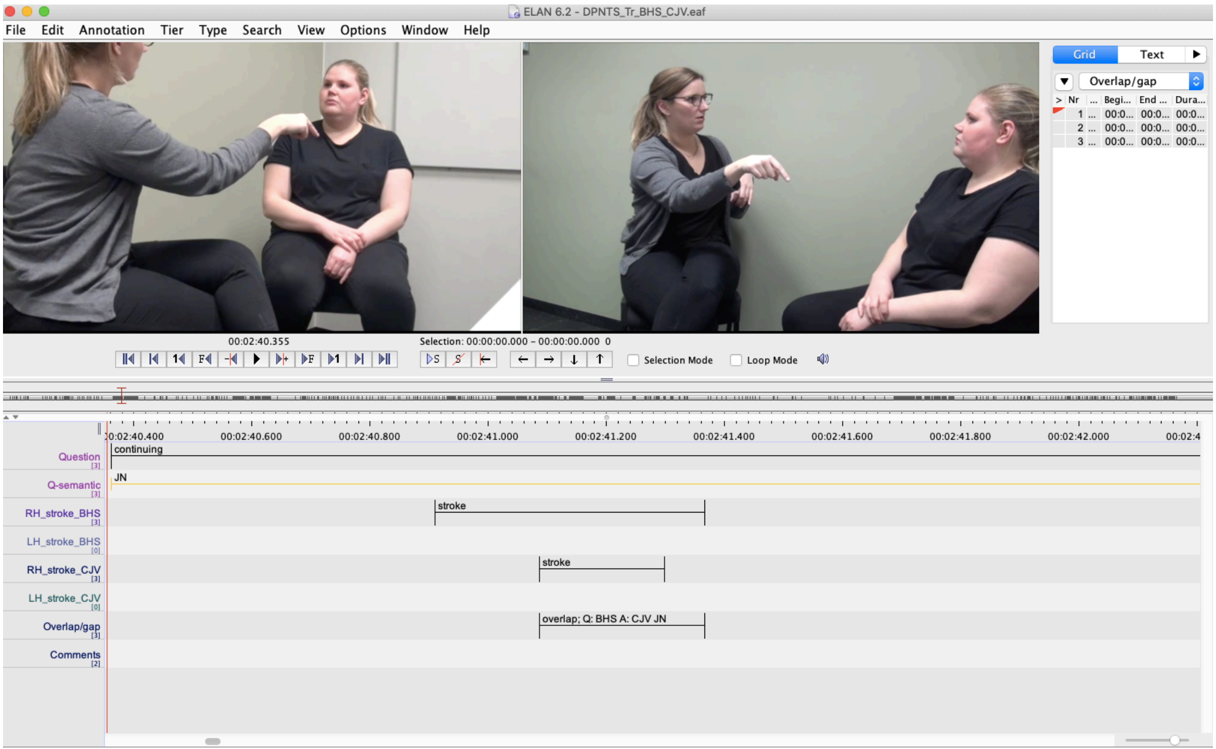
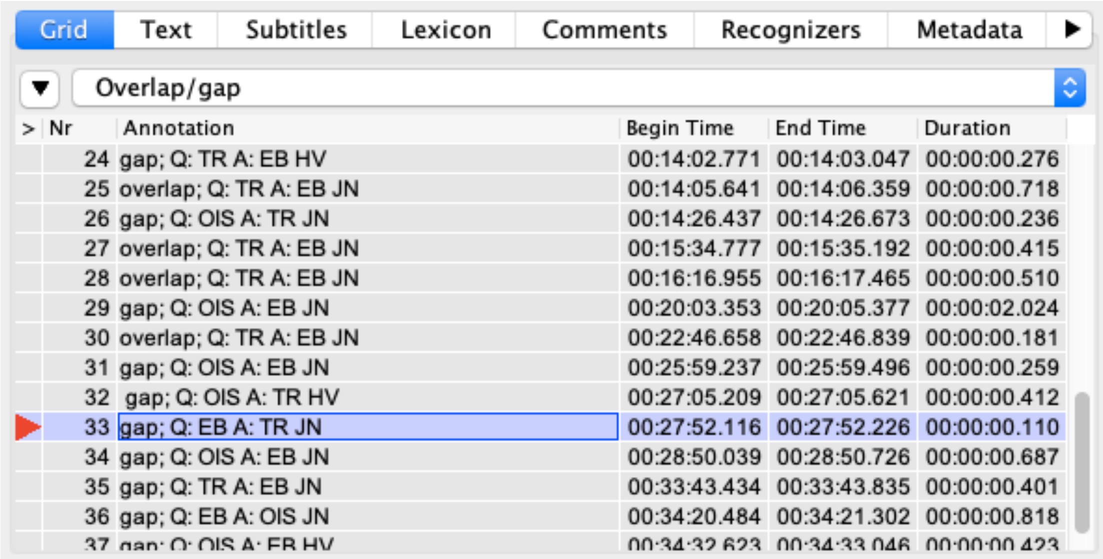


Figure 4.1: An example of an ELAN window from Group 1 with annotated strokes and overlap (Ferrara and Ringsø, 2021).

There are different types of tiers, independent tiers and referring tiers. Independent tiers hold annotations that are linked directly to a time interval, while referring tiers hold annotations that are linked to annotations on another tier (Brugman & Russel, 2009). The referring tiers are then linked to an annotation on a parent tier and does not necessarily have to be linked to a time interval, but if it is, the time interval is within the time interval of the annotation it is linked to on the parent tier (Brugman & Russel, 2009). In the corpus used in this study, both types of tiers occur. In Figure 4.1 the Q-semantic tier is a referring tier connected to the parent tier Question. However, only independent tiers were added during the data collection for this study.

It is also possible in ELAN, to search for and look through specific annotations in each file. This is done by selecting an annotation tier/category in the function, Grid or Text, in the top right area of the window (see Figure 4.1). If a specific tier is selected in Grid, a list of all annotations connected to this tier appears in a numbered list. The list includes the annotation name, begin time, end time, and duration (see Figure 4.2). If a specific annotation in the list is selected, the recording skips to this annotation in the workspace of the platform window. In the Text function, a list also appears, but only with the specific annotation names. With these functions, tiers/categories of annotations, or specific annotations are easily searched for, making the platform ideal to work on, trying to find generalizations of the language in question.



Nr	Annotation	Begin Time	End Time	Duration
24	gap; Q: TR A: EB HV	00:14:02.771	00:14:03.047	00:00:00.276
25	overlap; Q: TR A: EB JN	00:14:05.641	00:14:06.359	00:00:00.718
26	gap; Q: OIS A: TR JN	00:14:26.437	00:14:26.673	00:00:00.236
27	overlap; Q: TR A: EB JN	00:15:34.777	00:15:35.192	00:00:00.415
28	overlap; Q: TR A: EB JN	00:16:16.955	00:16:17.465	00:00:00.510
29	gap; Q: OIS A: EB JN	00:20:03.353	00:20:05.377	00:00:02.024
30	overlap; Q: TR A: EB JN	00:22:46.658	00:22:46.839	00:00:00.181
31	gap; Q: OIS A: EB JN	00:25:59.237	00:25:59.496	00:00:00.259
32	gap; Q: OIS A: TR HV	00:27:05.209	00:27:05.621	00:00:00.412
33	gap; Q: EB A: TR JN	00:27:52.116	00:27:52.226	00:00:00.110
34	gap; Q: OIS A: EB JN	00:28:50.039	00:28:50.726	00:00:00.687
35	gap; Q: TR A: EB JN	00:33:43.434	00:33:43.835	00:00:00.401
36	gap; Q: EB A: OIS JN	00:34:20.484	00:34:21.302	00:00:00.818
37	gap; Q: OIS A: EB HV	00:34:32.623	00:34:33.046	00:00:00.423

Figure 4.2: Example of what the Grid function looks like in ELAN.

The annotations created in ELAN can be exported in several different ways, depending on further use. In this project, annotations were exported as a CSV file to further examine the transition durations between the strokes of the signs in question-response sequences.

4.5 Movement phases and turn boundaries

The current study looks solely at manual signs. Following the phase movement coding scheme from Kita et al. (1998) which was discussed in subsection 3.2.1, this study treated each sign as potentially consisting of up to four movement phases, where the stroke is the phase that carries meaning and constitutes the turn boundary. In other words, the turn boundaries are measured from the last frame of the last stroke in the first signer's (the signer asking the questions) utterance to the first frame of the first stroke in the second signer's (the signer answering the question) utterance. The decision to measure stroke-to-stroke turn boundaries in this study was based on theories and findings from previous research. Findings that stated that the preparation phase of the sign corresponds to inbreaths in spoken language and that it is merely a preparation for speaking and not a part of the turn (McCleary & de Arantes Leite, 2012), p. 133). This theory is supported by de Vos et al. (2015) who found that only when the preparation phase, the hold, and the retraction phase were excluded from the turn in their analysis, the turn transition durations for NGT were in the same range as the cross-linguistic average for spoken languages measured by Stivers et al. (2009). In this study, therefore, there was only overlap if the last stroke of the question overlapped the first stroke of the answer. This decision was also made to make the current study comparable to the previous research mentioned above.

4.6 The annotation process

The data used in this study was drawn from nine video recordings from the NTS corpus. Six of the videos consisted of triads, and the other three consisted of dyads. Several relevant features had in each annotation file, already been annotated. These features included question type (yes/no, wh-, or alternative), whether the question had been asked manually or non-manually, and whether the question had been responded to.

Each participant had a separate ELAN file belonging to them, containing the above-mentioned information about the questions they had asked in the conversation. For each conversation, the annotation files per person were merged manually to create one annotation file per group, containing all the annotations relevant to questions for all participants in the conversations. Having all the information needed about each of the participants in the same file made the process of exporting the annotations easier, in the next stage of the study.

Before annotating the turn transition durations for the question-response sequences, new tiers, relevant to this study, were added to each merged file in ELAN. Each participant in the conversation group had two stroke tiers: right hand stroke per person and left-hand stroke per person, measuring the duration of their relevant strokes (either the last stroke of the last sign in the question or the first stroke of the first sign in the response). A tier for gap/overlap was added, measuring the duration of the potential gap or overlap between the stroke of the question and the stroke of the response, and a tier for comments, which included any comments the researcher had regarding any annotations (see Figure 4.3). In Table 4.3, a description of the tiers in conversation 1 is provided.

Table 4.3: Tier descriptions for Group 1.

Tier name	Definition
Question	Question
Q-manual	If the question was asked with manual elements
Q-non-manual	If the question was asked with non-manual elements
Q-semantic	Type of question (yes/no, content, alternative, or uncertain)
RH_stroke_BHS	Right hand stroke BHS
LH_stroke_BHS	Left hand stroke BHS
RH_stroke_CJV	Right hand stroke CJV
LH_stroke_CJV	Left hand stroke CJV
Overlap/gap	Overlap or gap
Comments	Any comments regarding the stroke or overlap/gap annotations

Though Table 4.3 describes the tiers from conversation group 1, the only tiers that are different in the other conversations are the stroke tiers. Each participant in each file had their own stroke tiers.

For each tier, there is also a number in brackets below the tier name. This number represents the total amount of annotations connected to that tier, in the annotation file. There are for instance 3 overlaps/gaps in the file in Figure 4.3.

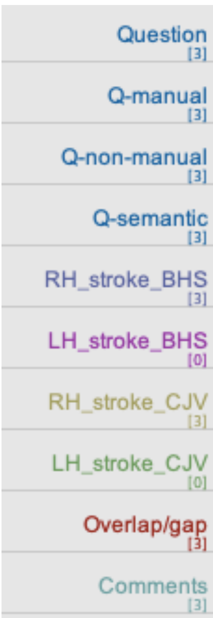


Figure 4.3. Screenshot of ELAN tiers from conversation group 1.

Once the tiers were added, the last stroke in the last sign of the question was annotated. Then the first stroke in the first sign of the answer to that question was annotated. All annotations were done manually for each question. The strokes were measured from the first frame where the location, hand shape, and orientation of the sign were in place until the first frame where the location, hand shape, or orientation showed change or movement, or the sign transitioned into another sign or into the retraction phase (see Figure 4.4 and Figure 4.5).

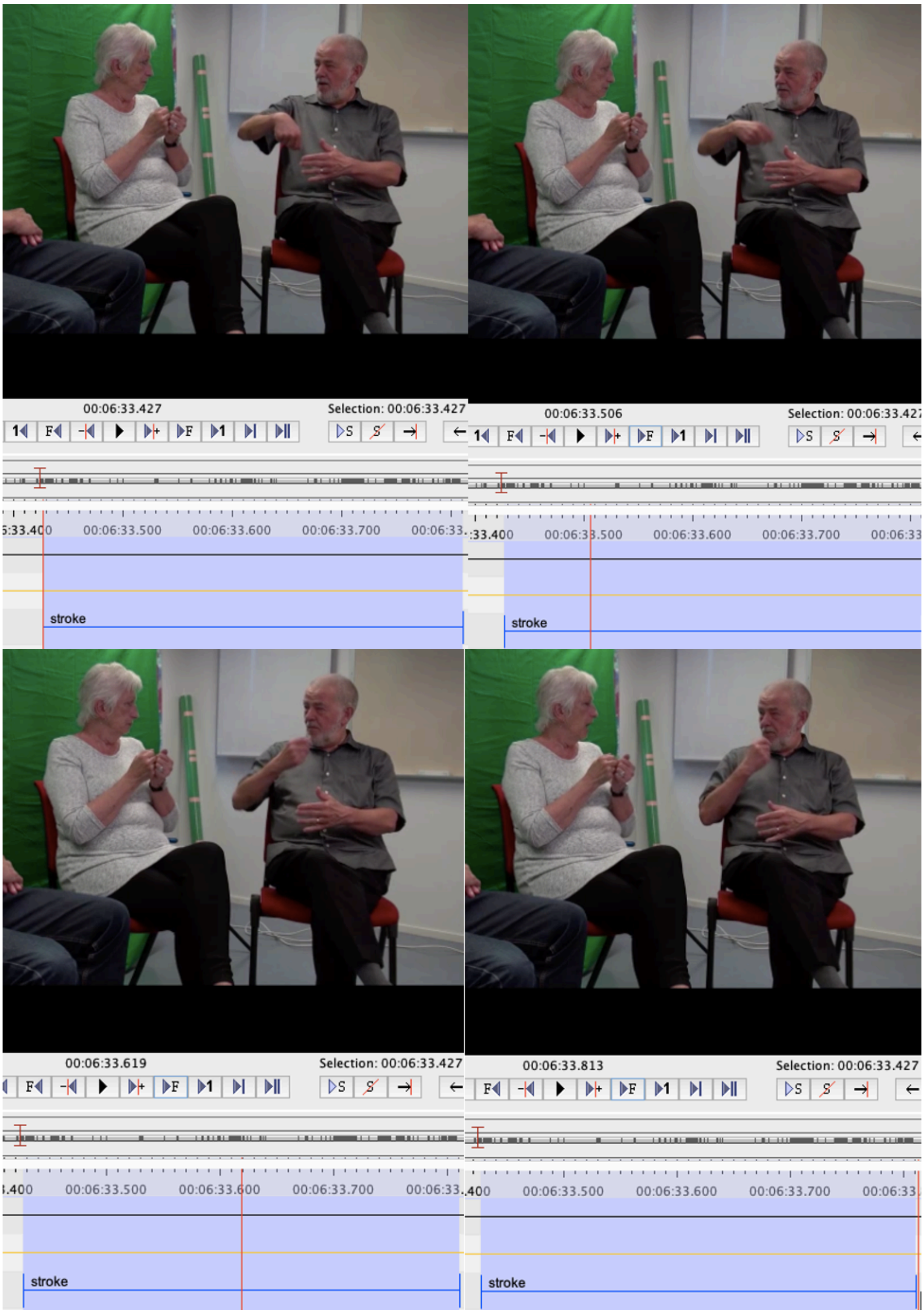


Figure 4.4: The stroke phase of a question asked by OIS (the man on the right) in conversation 7 (Ferrara and Bø, 2015).

Figure 4.4 illustrates four different frames during the stroke phase of the last sign in a question, from when the stroke starts until the last frame of the stroke, before the retraction phase begins.

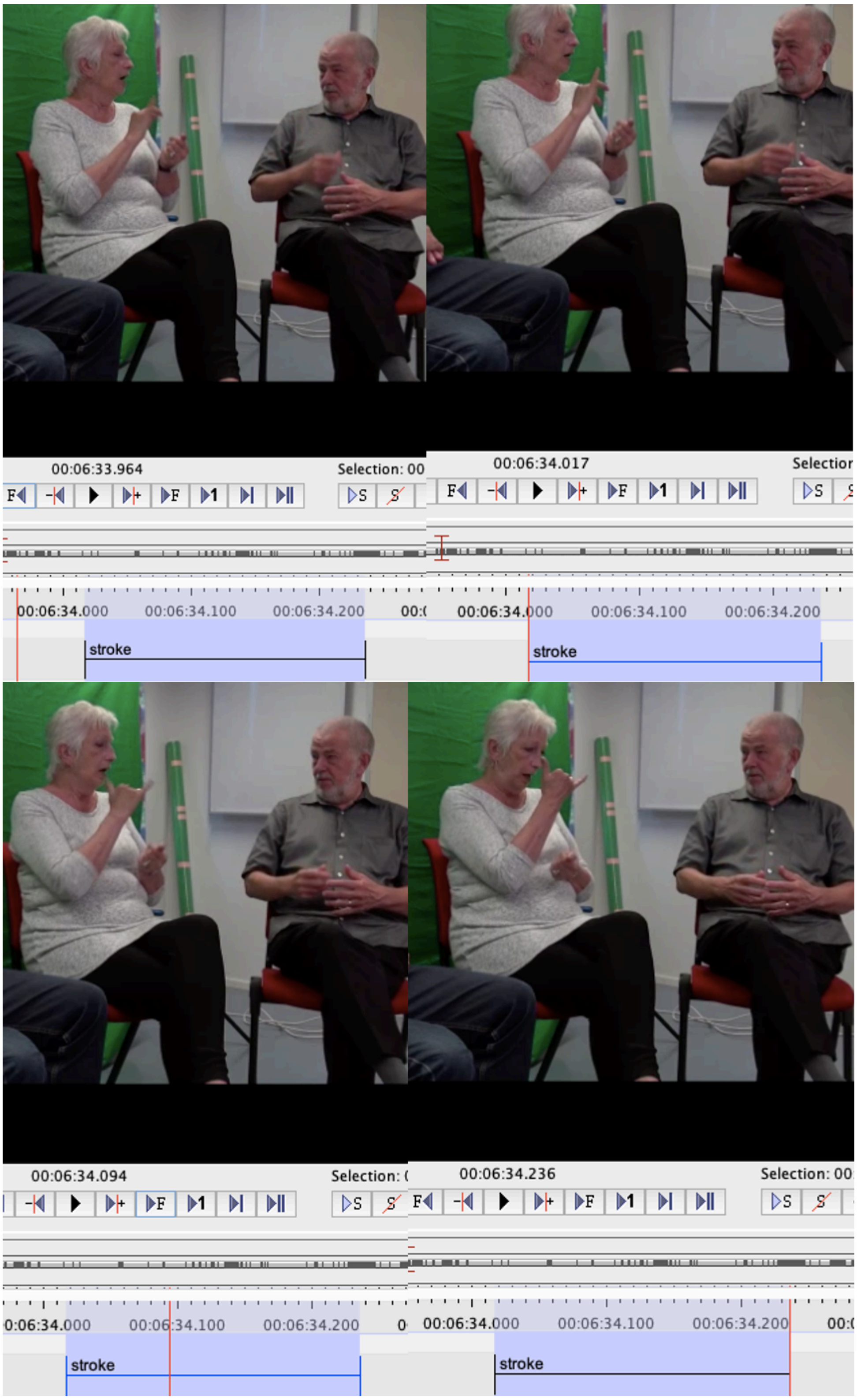


Figure 4.5: The stroke phase of EB's (the woman in the middle) answer to the question asked by OIS (on the right) in Figure 4.4 (Ferrara and Bø, 2015).

Figure 4.5 illustrates four different frames from the stroke phase of the answer to the question asked in Figure 4.4. In the first frame EB has not quite gotten her hand into the sign's position, and the stroke is therefore annotated starting in the second frame of the illustration. It ends in the fourth frame after the movement of the sign is finished, before retraction.

After these annotations were in place, the duration of the gap or overlap between these strokes was annotated. This was done by marking the area from the end of the final question stroke to the beginning of the first answer stroke. The gap and overlap annotations included information about whether it is a gap or an overlap, the type of question that was asked, as well as information about which signer asked the question, and which signer answered (see Figure 4.6). This process was repeated in all nine conversation group files.

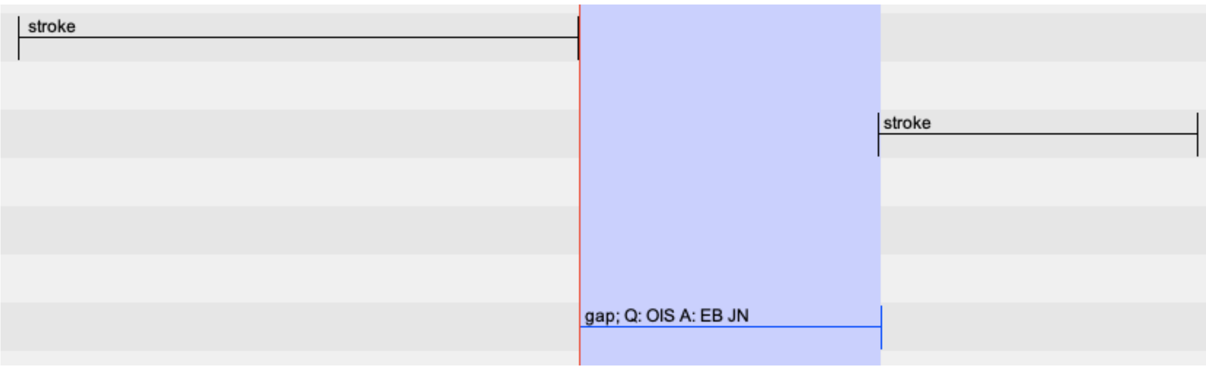


Figure 4.6: The gap between OIS's question (Figure 4.4) and EB's answer (Figure 4.5).

Figure 4.7: An example of an overlap of two strokes in the conversation in Group 7.

Figure 4.7 illustrates an instance of overlapping signing between TR and EB, where EB started signing before TR got to the last stroke of the question.

As mentioned at the beginning of this subsection, the questions had already been annotated regardless of any response. While going through the questions for this study, all questions that had no response were automatically excluded from the data set and not annotated. In addition to this, all questions that were asked non-manually were excluded from the data set because this study solely looked at transition durations between manual signs. A third portion of questions were excluded due to unclear contexts surrounding transitions. All questions were examined three times, and these were transitions that were still unclear after the third round of annotating.

4.6.1 Difficult cases

Decisions had to be made regarding signs such as pointing, palm up, as well as repetitions of the stroke phase in certain signs. Pointing and palm up signs often include a natural hold. This may occur while the signer asking the question is waiting for the response of the second signer. In such signs, it is difficult to distinguish between the stroke and hold because there is no visible transition between the two phases. However, if, following the previously mentioned coding scheme, the stroke phase is over once the sign has achieved its hand shape, location, orientation, and movement, then this should apply to signs such as pointing and palm up as well.

The transition from the sign to the hold in pointing usually only consisted of a change in movement (see Figure 4.8 and Figure 4.9). Though it was harder to determine exactly when this change happened, it was possible. Pointing signs were therefore annotated as strokes until the first frame where movement ceased.

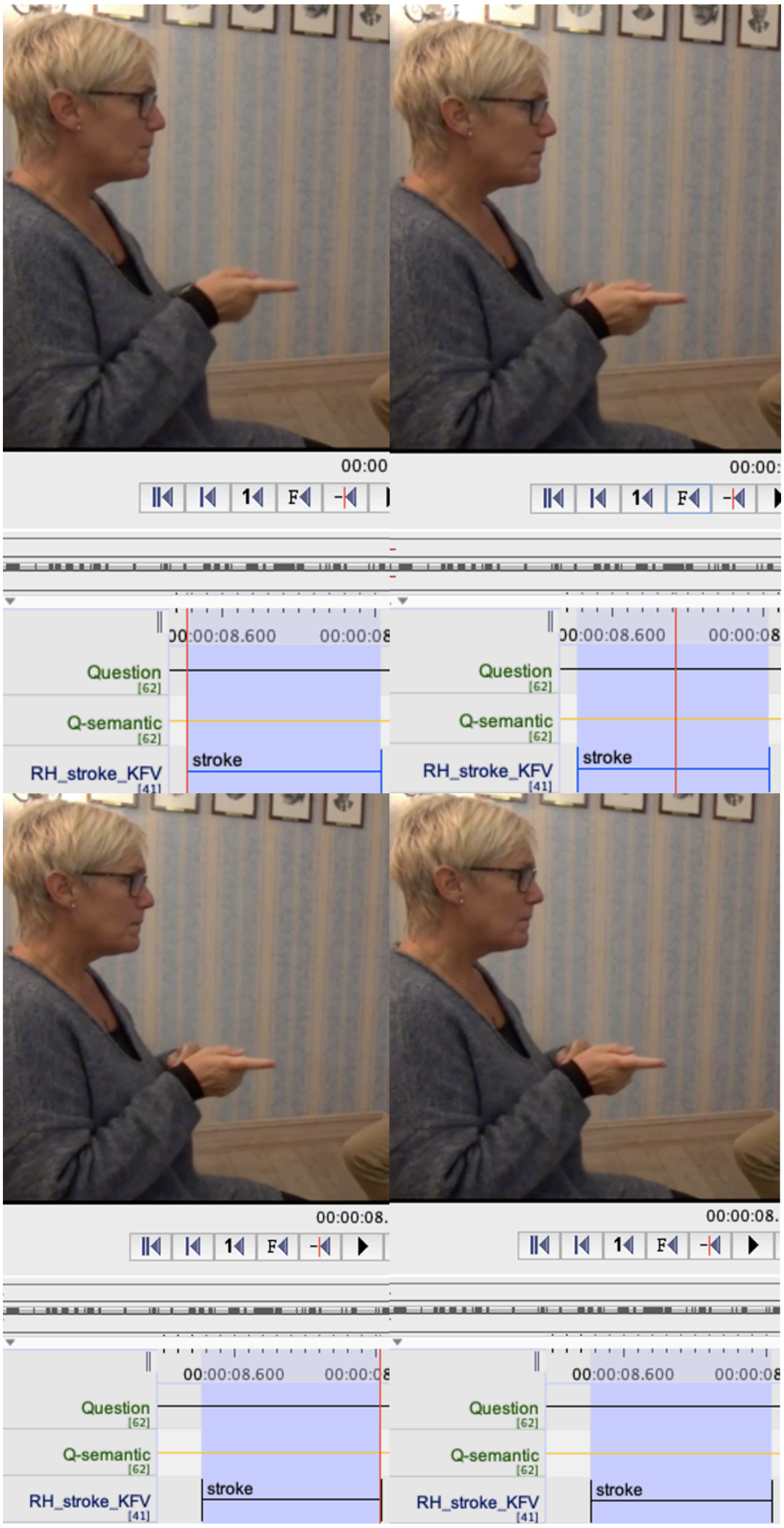


Figure 4.8: Example from Group 6 of a pointing sign as the last stroke in a question (Ferrara and Ringsø, 2021).

As illustrated in Figure 4.8, there are only subtle differences in the frames of EB's last stroke. In the first frame illustrated, her right hand covers her left hand, but as the stroke progresses, her left hand is showing more and more, indicating downward movement of the right hand.

However, in the last two frames, there is not a noticeable difference anymore, indicating that movement of the right hand has stopped in the penultimate frame and that the last frame is part of the hold rather than the stroke.

The palm up sign, on the other hand, did not always consist of movement once the hand shape, orientation, and location were achieved and was therefore slightly more challenging to annotate the stroke for. Looking closely, however, some slight movement in the shoulders or elbows could be seen, as if the signer relaxes those limbs slightly more during the hold than during the stroke. The stroke was therefore measured from the first frame where the hand shape, location, and orientation were achieved until such a movement could be seen. In cases where this movement could not be detected, the question was excluded from the study due to the importance of consistency in annotating.

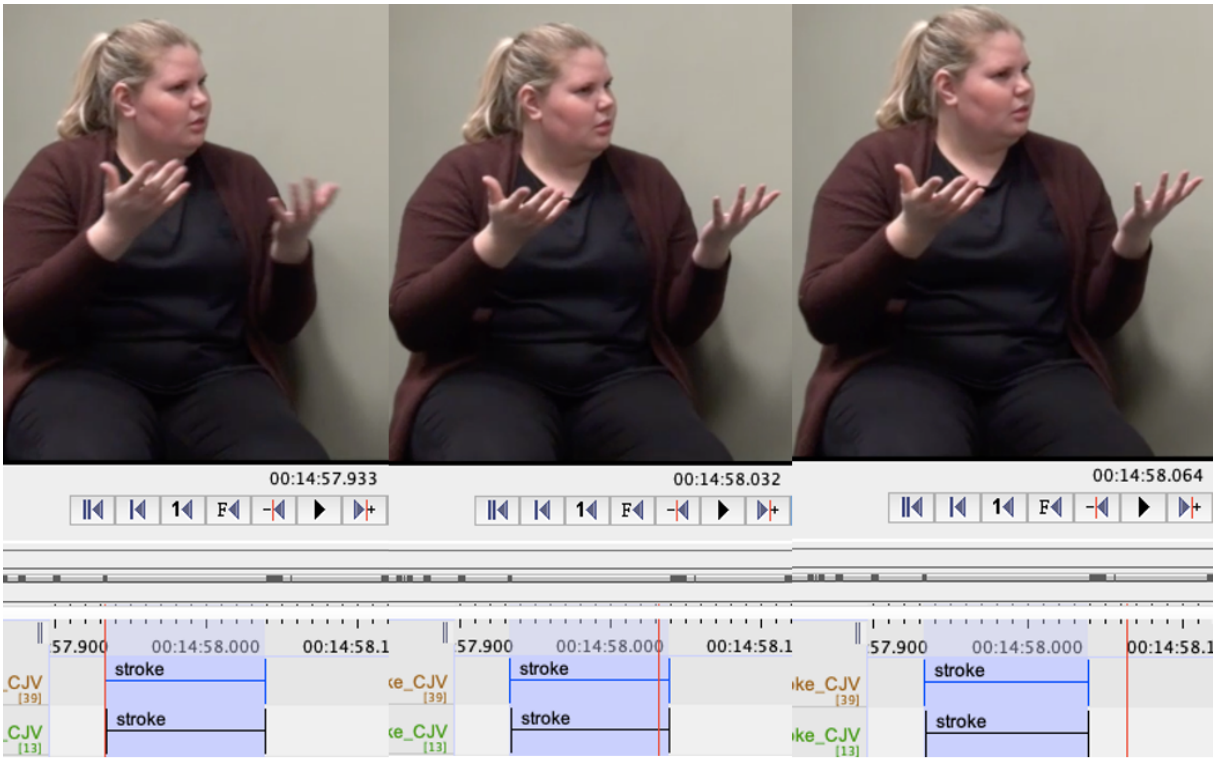


Figure 4.9: Example from Group 2 of a palm up sign as the last stroke in the question CJV is asking (Ferrara and Ringsø, 2021).

Figure 4.9 illustrates the palm up sign in a question-answer sequence between CJV and TJ. There is outward movement of the hands from the first to the second frame, but in the third frame, where the stroke has ended, the movement has also stopped indicating that the third frame is part of the hold rather than the stroke.

In some cases, the stroke phase of a sign was repeated several times. This seemed, in some cases, to serve the same function as a hold. In these cases, the stroke would be repeated until the next signer responded. In other cases, the repetition served as an emphasizing element. If the second signer enthusiastically responds positively to the question asked, the stroke phase sign for 'yes' could be repeated several times for instance. In such cases as these two, the decision was made to only measure the first stroke. This is because, even if the repetitions add some sort of extra-linguistic element to the utterance, the meaningful part of the sign has already been conveyed by the first stroke.

Figure 4.10: Example from Group 7 of the repetition of a sign (by the man on the left) (Ferrara and Bø, 2015).

Figure 4.10 illustrates four frames from a stroke by TR where he signs YES, and the movement of the sign is completed by the fourth frame. In Figure 4.11, however, where four frames following the stroke are illustrated, he has not preceded to the retraction phase, but has repeated the movement of the sign YES.

Figure 4.11: Example from Group 7 of the repetition of a sign (by the man on the left) (Ferrara and Bø, 2015).

4.7 The data set

When all the relevant annotations were added to the conversation group files, the final data set consisted of 159 annotated questions, with 100 gap transitions and 59 overlap transitions. The final dataset consisted of 54 wh-questions, 89 yes/no questions, 11 alternative questions, and five questions where the type was unclear.

Three conversation groups contained the majority of the questions (84.9%). These groups were group 2, which contained 40 questions, with 18 overlaps and 22 gaps, group 6, which contained 57 questions, with 22 overlaps and 35 gaps, and group 7, which contained 38 questions, with 11 overlaps and 27 gaps (see Table 4.4).

Table 4.4: Distribution of transitions in the dataset across question types and the gap vs. overlap categories, for the 9 groups.

Group	Questions	Yes/no	Wh-	ALT	Unclear	Overlaps	Gaps
1	3	3	0	0	0	1	2
2	40	12	22	5	1	18	22
3	6	6	0	0	0	1	5
4	3	2	1	0	0	2	1
5	3	1	2	0	0	0	3
6	57	39	12	4	2	22	35
7	38	21	14	1	2	11	27
8	6	4	1	1	0	2	4
9	3	1	2	0	0	2	1
Total	159	89	54	11	5	59	100

4.8 Inter-rater reliability agreement

To test the reliability of the main rater of the overlap/gap transition annotations, a second rater annotated 20 percent of the data set.

Two types of quantitative analysis of inter-rater agreement were then done. First, the agreement in classification was assessed (gap vs. overlap) and apply the Cohen's kappa to calculate agreement by chance (Cohen, 1960). The Cohen's kappa was equal to 0.77, which can be classified as moderate to strong agreement, and it was significantly different from chance agreement ($p<0.001$).

Second, because duration of the gaps and overlaps was a crucial measure for the analysis, the absolute difference for the annotations by the two raters was also calculated. The median difference was 45 ms, which corresponded to 1.3 frames based on the 30fps framerate, and the mean difference was 146 ms, which corresponded to 4.4 frames. The mean was much higher than the median due to a small number of outliers with a large absolute difference between the two raters.

On the one hand, the two raters generally agreed on the gap vs. overlap label, and the duration of the annotations were well aligned. However, there were also some cases of clear disagreement for both label assignment and duration annotation. These cases were carefully analyzed and discussed before a final round of annotations was done by the first rater.

4.9 Method of analysis

To find out whether transition durations in NTS question-answer sequences are in the range observed for other spoken and signed languages, a quantitative analysis was done in a statistics software with the help of Vadim Kimmelman.

All the Overlap/gap annotations were exported from ELAN and examined R version 4.1.3 (R Core Team, 2022) and R Studio version 2022.02.1 (RStudio Team, 2022), where the transition durations were investigated. For every question-answer sequence, turn transition duration was represented with the Floor Transfer Offset (FTO) used in De Reuter et al. (2006) and defined as “the difference between the time that turn starts and the and the moment the previous turn ends” (p. 516). This value is given in positive milliseconds when the transition is a gap, and negative milliseconds when the transition is an overlap.

The mean and median durations for the whole data set were calculated. The differences in duration depending on which participant asked or answered the question were also calculated in R Studio. To investigate the effect the age range of the participants and the question types had on the average transition duration, a mixed effects linear regression was built using R Studio packages to calculate if there was a statistical difference.

After the quantitative analysis was completed, to further examine specific examples of transitions that were particularly long, a qualitative analysis was done by observing the examples in ELAN and describing the actions surrounding the overlaps and gaps to investigate whether any explanation could be offered.

5 Analysis and results

The current research aims to investigate whether turn transition durations in NTS question-answer sequences are in the range observed for other spoken and signed languages, as well as whether the average durations are variable between individuals and if it is affected by either the type of question or the age of the signers. The data set contained 159 transitions, with 100 gaps, and 59 overlaps. In this chapter, the results from the statistical analysis of this data set are presented, as well as a qualitative analysis, investigating a selection of transitions that did not fall within the previously mentioned cross-linguistic range.

5.1 Quantitative analysis and results

When stroke-to-stroke turn boundaries were considered, the mean duration of transitions for the full data set was 185 ms, and the median was 138 ms. The standard error of the mean for the duration was 48, and the 95% Confidence Interval of the mean was [91, 280]. The inter-quartile range for the data set was from -137 ms to 518 ms, and the full range was from -1468 ms to 2024 ms.

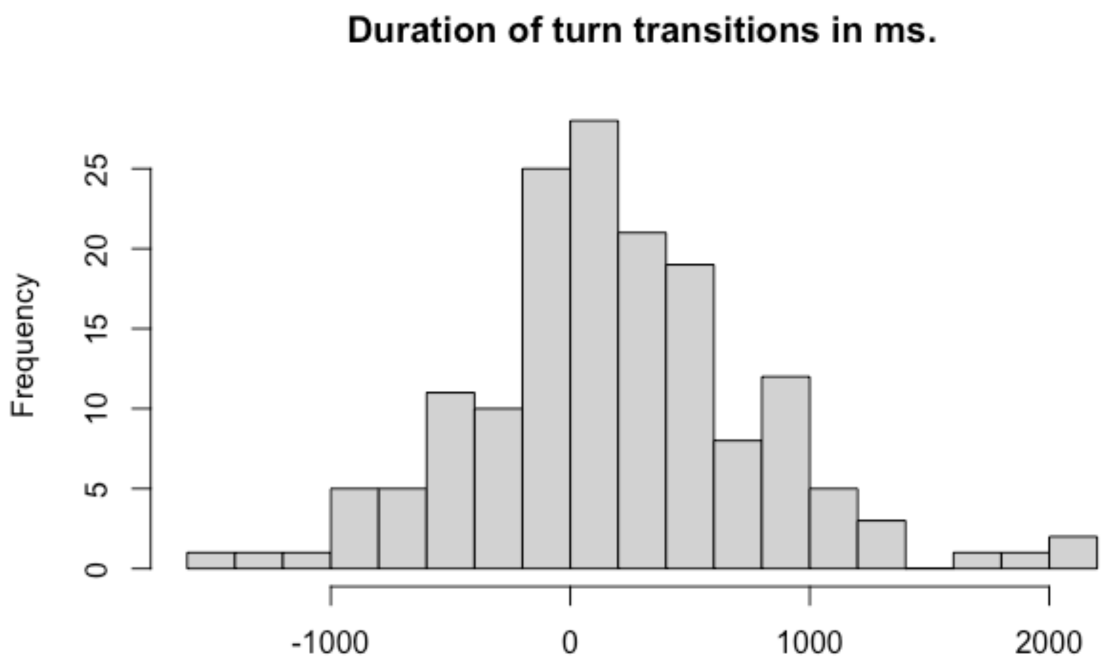


Figure 5.1: Distribution of transition duration in ms for the full data set.

Figure 5.1 shows the distribution of transition durations in ms for the full data set. Most transitions were very close to 0, on either side, i.e., most transition durations, both gaps and overlaps were very short. It also shows that there were outliers on both sides, i.e., both a few long gaps, and a few long overlaps. The longest gap was 2024 ms, and the longest overlap was 1468 ms (more information on this in subsection 5.2).

5.1.1 Individual differences

In the data set, there were 19 different participants. 15 of the participants asked questions, and 13 answered questions. Individual differences in transition durations occurred depending on who was asking the question, and who was answering the question (see Fig. 5.3 and Fig. 5.4), but as can be seen in Table 5.1 and Table 5.2, the distribution of questions and answers per participant was not balanced.

Table 5.1: Distribution of questions asked per participant.

Participant	**Questions asked**
BHS	3
CJV	38
EB	4
EL	3
EMN	3
ER	3
IMH	3
KFV	37
MF	2
MS	16
OIS	24
ØSR	4
PN	6
TJ	2
TR	11

Table 5.1 shows that most of the questions in the data set were asked by only a few signers (CJV, KFV and OIS), and that most of the signers who asked questions, asked less than five each.

Table 5.2: The distribution of questions answered per participant.

Participant	Questions answered
CJV	10
EB	30
KFV	6
MF	1
MS	21
OIS	2
ØSR	30
TH	3
TJ	38
TR	6
TR2	4
PS	3
LPL	5

A similar distribution can be seen in Table 5.2 which shows the signers who answered questions, and the number of questions each signer answered. This too is an unbalanced distribution where a few signers represent the majority of questions answered.

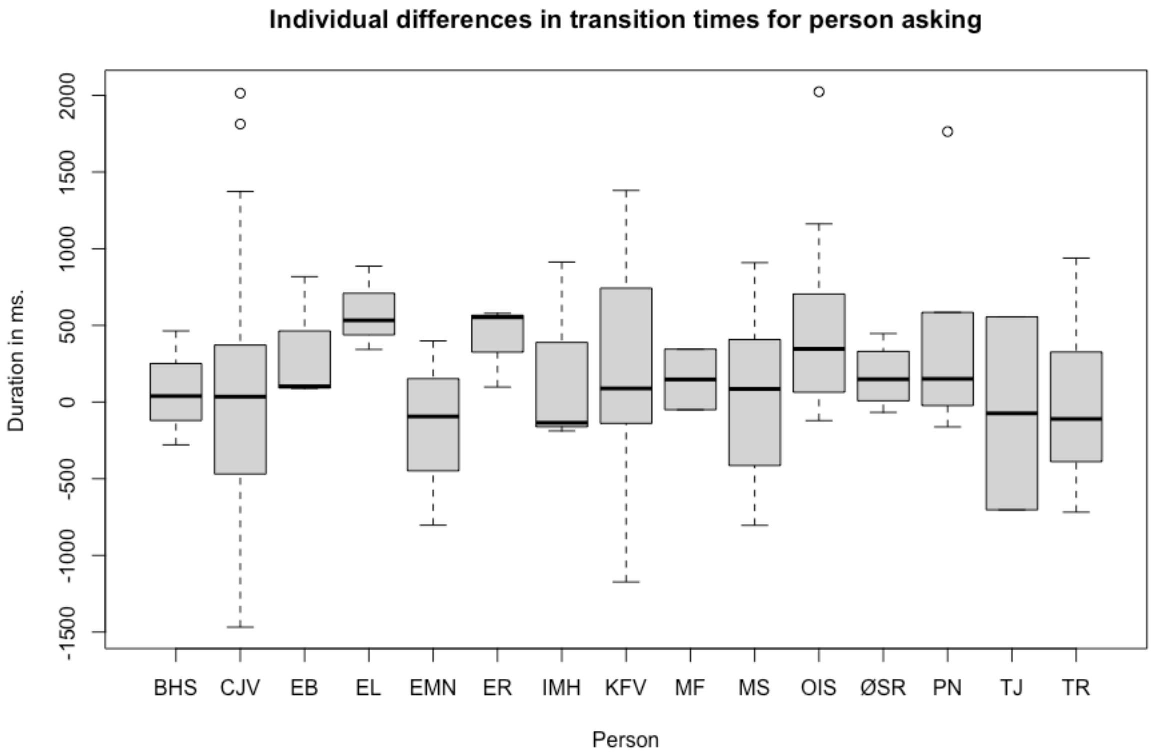


Figure 5.2: Individual differences in transition durations for person asking questions.

As mentioned, at the start of this subsection, the number of questions asked per person varied. Only five individuals asked more than ten questions, CJV, KFV, MS, OIS, and TR, with 38, 37, 16, 24, and 11 questions respectively. As can be seen in Figure 5.2, the signers who asked the most questions, generally had the greatest range in durations. The two signers who asked the most questions (CJV and KFV) show the greatest range, though CJV asked questions producing both longer gaps and longer overlaps than KFV. CJV had a slightly negatively skewed distribution, and KFV a positively skewed distribution. Figure 5.2 also shows that MS and TR had a similar distribution of durations, but MS had a negatively skewed distribution, whereas TR had a positively skewed distribution. OIS had a slightly positively skewed distribution, and when he asked questions, it resulted in mostly gaps, with less than 25 percent overlaps.

The remaining signers show varying distributions of transition durations, but as they all asked less than ten questions, and a few of them less than five, they will not be discussed further.

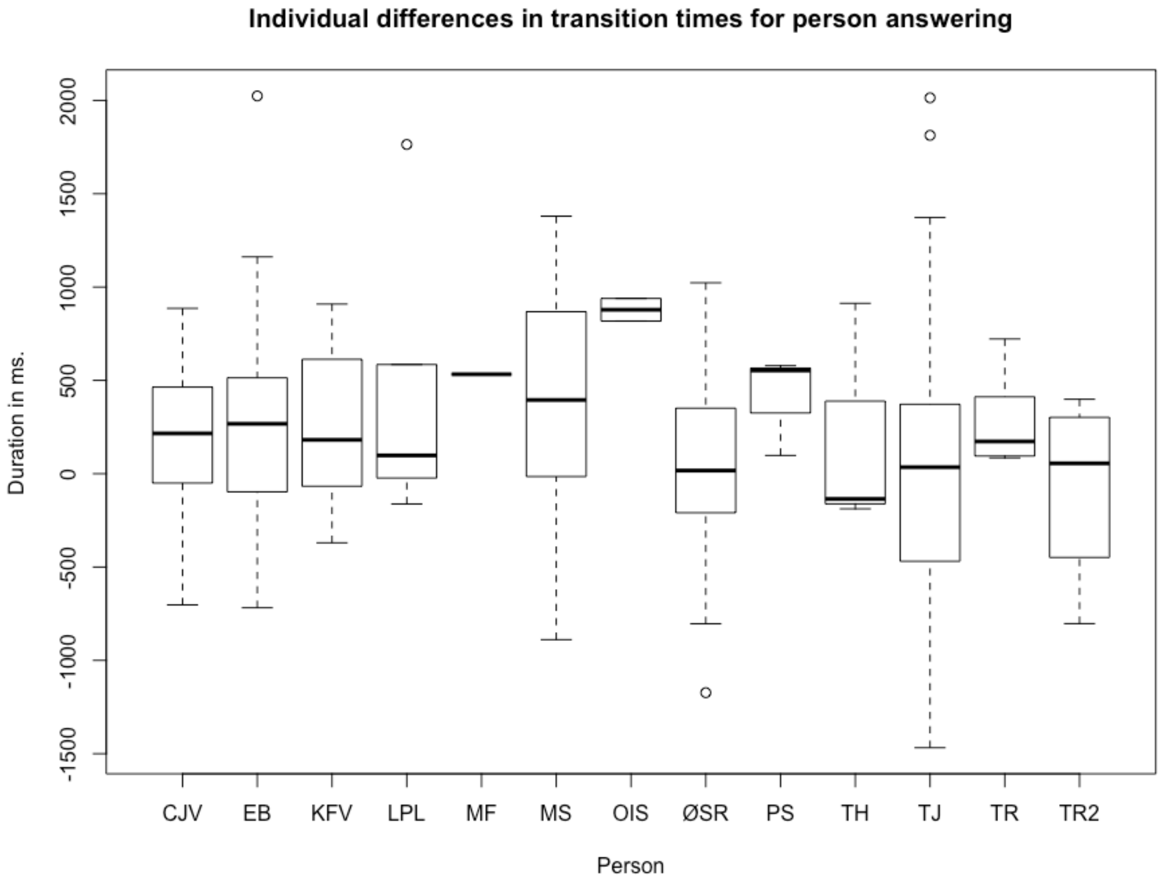


Figure 5.3: Individual differences in transition durations for person answering questions.

Figure 5.3 shows the individual differences in transition durations per person answering questions. Just as with the number of questions asked by each signer, the number of questions each signer answered also varied greatly. Four signers answered more than ten questions, EB, MS, ØSR, TJ, with 30, 21, 30, 38 questions respectively. Transitions preceding TJ's answers varied most in duration, but he was also the signer answering the most questions. TJ's distribution was negatively skewed with a median close to 0 ms, while ØSR's distribution was positively skewed, with a median close to 0 ms. MS and EB showed a similar durations range, with almost 75 percent of their data points above 0 ms, i.e., gaps. EB's distribution is negatively skewed however, and MS's distribution, slightly positively skewed.

5.1.2 The effect of question type and age

In the full data set, there were 159 questions, 11 alternative questions, 54 wh-questions, 89 yes/no-questions, and five cases where the type of question was unclear. The five unclear cases were excluded from the analysis of question types.

The mean transition duration for alternative questions was -75 ms, 217 ms for wh-questions, and 186 ms for yes/no-questions. The distribution is shown in a boxplot in Figure. 5.4.

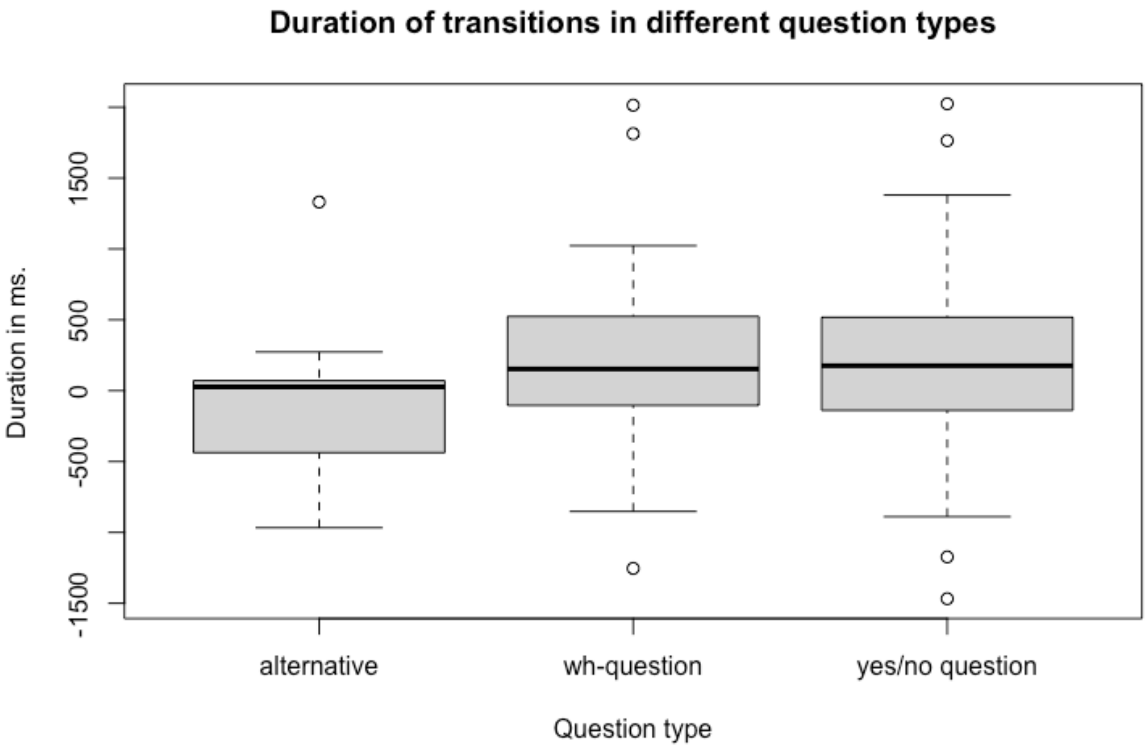


Figure 5.4: Distribution of transition duration in 3 question types.

Looking at the boxplot in Figure 5.4, wh-questions and yes/no questions seem to be very close, whereas alternative questions are different in that they show a negative mean transition time. However, there were only 11 alternative questions.

The effect of the age-range of participants asking or answering the question on the average transition durations was also investigated. The number of questions asked and answered by each age group was not balanced, however, as can be seen in Table 5.3 and Table 5.4. The distribution of transition durations per age range asking and answering questions are shown in Figure 5.5 and Figure 5.6.

Table 5.3: Distribution of questions asked per age range.

Participant age range	Questions asked
18-29	54
30-40	7
40-50	19
50-60	40
60-70	24
70-80	15

As shown in Table 5.3, the questions asked per age range varied, and because the age range of 30-40 only asked 7 questions, it will not be further discussed.

Table 5.4: Distribution of questions answered per age range.

Participant age range	Questions answered
18-29	57
30-40	37
40-50	21
50-60	6
60-70	2
70-80	36

Table 5.4 shows that, as with questions asked per age range, the number of questions answered per age range also varies, and because age ranges 50-60 and 60-70 answered less than ten questions each, they will not be further discussed.

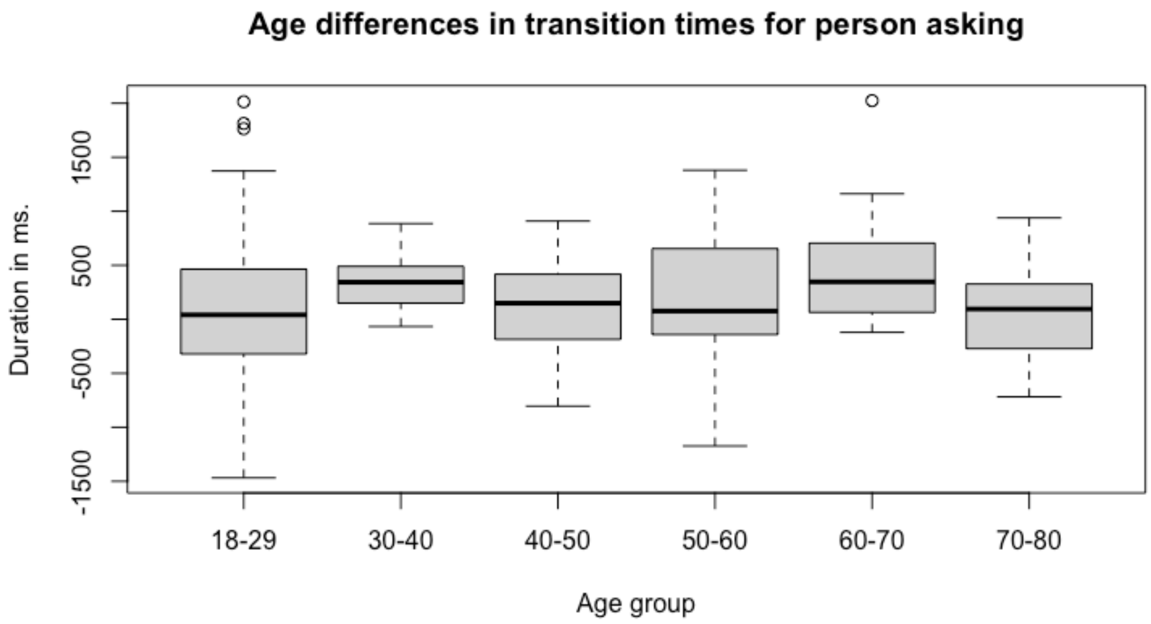


Figure 5.5: Distribution of transition durations per age range asking questions.

Figure 5.5 shows the differences in transition durations per age range asking the question. Observationally, there does not seem to be great differences between the groups, though only two age ranges include outliers, all of them gaps.

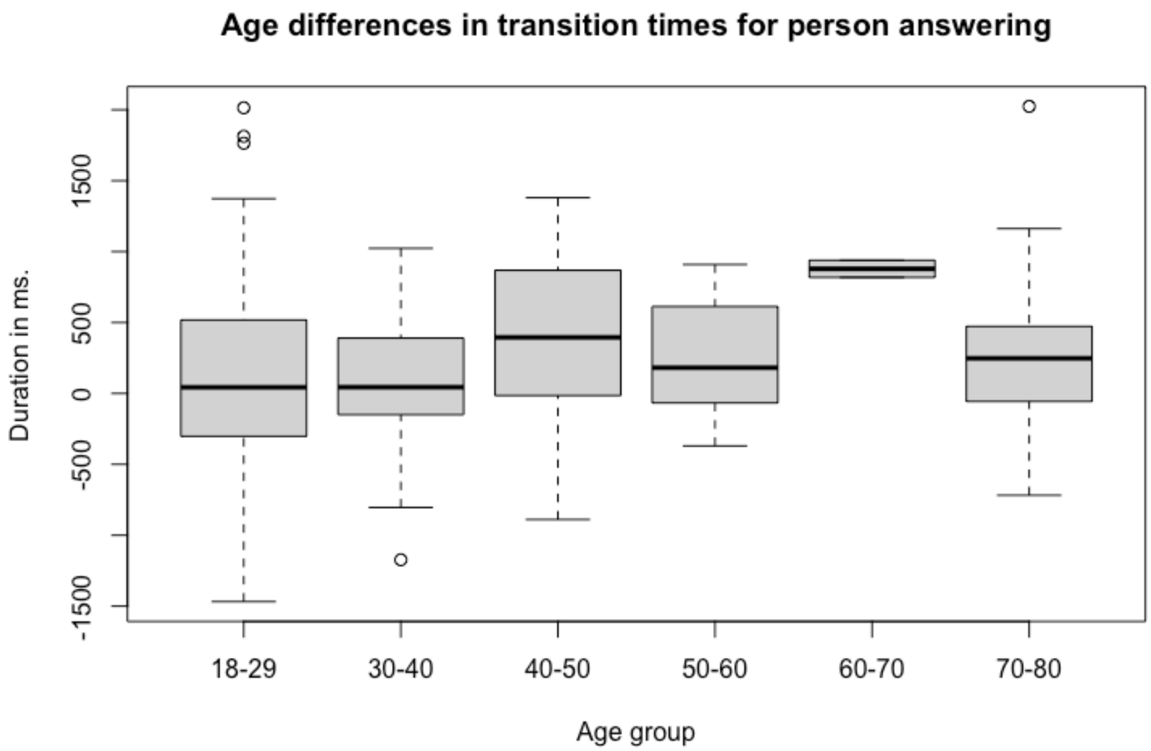


Figure 5.6: Distribution of transition durations per age range answering questions.

Figure 5.6 illustrates differences in durations per age range answering the question and there is not, observationally, a great difference between the ranges. Age range 18-29 shows the greatest range which is in line with it being the range with the most data points. The distribution of age ranges 18-29 and 30-40 are close to normal, but slightly positively skewed, with a median close to 0 ms, i.e., there are more gaps than overlaps. Age range 40-50, on the other hand, had a median closer to 500 ms, but transitions preceding answers from this range were also mostly gaps.

To check whether the question types or age-range of the participants were statistically different, a mixed effects linear regression was built using l*me4* (Bates et al., 2015) and *lmerTest* (Kuznetsova et al., 2017) packages, with duration as the dependent variable and question type and age-range of the person asking or answering the question as the independent variables, with random intercepts per participant asking and answering the questions.[1] Because question type had three levels, orthogonal contrast coding was used (comparing wh-questions to yes/no questions, and comparing alternative questions to the average of the other two categories). The model showed that the differences in duration between different question types or age ranges do not reach statistical significance. Statistical evidence for the effect of question type or age range of duration was thus not found.

5.2 Qualitative analysis and results

Because there is an expectation for minimal gap and overlap based on previous research, a qualitative analysis was done to investigate whether the instances of long gaps and overlaps shared any common characteristics, or if a reason for the extended duration could be observed. In the following two subsections, these transitions are described one by one. Only transitions of gaps and overlaps with a duration of 1000 ms or more were further investigated which included a total of 12 gaps and 3 overlaps (See Table 5.5).

[1] A model with random slopes per participant for question type effect does not converge due to a singular fit. A model with interaction between age and question type does not converge for the same reason.

Table 5.5: All individual overlaps and gaps from the data set exceeding 1000 ms.

Participant asking (age range)	Participant answering (age range)	Type of question	Duration in ms
KFV (50-60)	MS (40-50)	Yes/no	1110
KFV (50-60)	ØSR (30-40)	WH	1023
KFV (50-60)	MS (40-50)	Yes/no	1075
KFV (50-60)	MS (40-50)	Yes/no	1380
KFV (50-60)	MS (40-50)	Uncertain	1073
PN (18-29)	LPL (18-29)	Yes/no	1764
CJV (18-29)	TJ (18-29)	Yes/no	1373
CJV (18-29)	TJ (18-29)	ALT	1331
CJV (18-29)	TJ (18-29)	WH	2014
CJV (18-29)	TJ (18-29)	WH	1813
OIS (60-70)	EB (70-80)	Uncertain	1162
OIS (60-70)	EB (70-80)	Yes/no	2024
KFV (50-60)	ØSR (30-40)	Yes/no	-1173
CJV (18-29)	TJ (18-29)	Yes/no	-1468
CJV (18-29)	TJ (18-29)	WH	-1254

5.2.1 Gaps

Four conversation groups experienced gaps exceeding 1000 ms in their interactions. Group 2, Group 6, Group 7, and Group 8.

In the conversation between KFV, MS, and ØSR (Group 6), there were five such gaps, four transitions that followed a question asked by KFV and answered by MS, and question that was asked by KFV and answered by ØSR. In one instance, where KFV asks a question, and MS answers, MS shakes his head slightly and flattens his lips before replying manually, as shown in Figure 5.7.

Figure 5.7: Snapshots of a long gap in conversation 6 between KFV's (on the left) question and MS's (on the right) answer (Ferrara and Ringsø, 2021).

In the four other instances, the person answering the question shows signs of thinking before replying. This was done either by mouthing 'uhm' (Figure 5.8), an upward gaze (Figure 5.9), or an upward gaze combined with a touch of the chin (Figure 5.10).

Figure 5.8: Snapshots of a long gap from conversation 6 between KFV's (woman in the first frame) question and ØSR's (man on the right) answer (Ferrara and Ringsø, 2021).

Figure 5.8 shows KFV asking the question (top two pictures), and ØSR mouths something similar to 'uhm', seemingly thinking before beginning to reply to the question.

Figure 5.9: Snapshots of a long gap in conversation 6 illustrating MS's (man on the right) upward gaze while thinking (Ferrara and Ringsø, 2021).

Figure 5.9 shows KFV asking the question, and holding the sign while MS gazes upwards, thinking for a moment before beginning to reply.

Figure 5.10: Snapshots of a long gap in conversation 6 illustrating MS's (man on the right) upward gaze and chin touch (Ferrara and Ringsø, 2021).

Figure 5.10 shows KFV asking the question and retracting, while MS gazes upwards and puts his index finger to his chin, thinking, before beginning to reply.

In the conversation between PN, LPL, and TR2, there was one instance of a gap transition with a duration exceeding 1000 ms. The transition followed a question asked by PN which was answered by LPL. As shown in Figure 5.11, before responding, LPL sits very still, with a fixed gaze, seemingly thinking, and does not reply until PN starts signing again.

Figure 5.11: Snapshots from conversation 8 illustrating LPL (on the left) sitting still, not responding right away (Ferrara and Ringsø, 2021).

In the conversation between CJV, and TJ (Group 2), there were four instance of gap transitions with durations exceeding 1000 ms. All the transitions followed questions asked by CJV which were answered by TJ. In the following illustration (Figure 5.12), CJV asks a question, and the gap lasts for 1373 ms before TJ replies. During the gap TJ stays still for a moment, holding his hands still in the signing space, then nods his head backwards a little while opening his mouth, as in understanding the question, then he begins to reply manually.

Figure 5.12: Snapshots from a long gap in conversation 2 illustrating TJ (on the right) staying still with his hands in the signing space while thinking before replying to CJV's question (Ferrara and Ringsø, 2021).

Two other gaps from their conversation, that lasted 1331 ms and 1813 ms, were very similar to the one illustrated in Figure 5.12. TJ has his arms in the signing space, staying still, with an unfocused gaze before beginning to reply. The last instance lasted 2014 ms and was one of the only two gaps in the whole data set that exceeded 2000 ms. As shown in Figure 5.13, TJ goes into resting position, blows air into his cheeks while gazing upwards, as if the question is hard to answer.

Figure 5.13: Snapshots of the longest gap from conversation 2, illustrating TJ's (on the right) upward gaze and filling his mouth with air before replying (Ferrara and Ringsø, 2021).

In the conversation between OIS, and EB, and TR (Group 7), there were two instance of gap transitions with durations exceeding 1000 ms. Both the transitions followed a question asked by OIS which was answered by EB. Though it is not very clear in the illustration below (Figure 5.14), EB (in the middle) takes a breath during the 1162 ms long gap, before beginning to reply to OIS (on the right).

Figure 5.14: Snapshots of a long gap in conversation 7 illustrating KFV (woman in the middle) taking a breath while thinking before replying to OIS's (man on the right) question (Ferrara and Bø, 2015).

In the second instance of a long gap in that conversation, EB gazes upwards, thinking, before replying to OIS.

In 11 out of 12 gaps, the person responding to the question seemed to be thinking before replying which resulted in a longer than usual gap. This was shown by participants having an unfocused upward or sideway gaze, or by taking a breath, or just staying still. In one case, there was a non-manual response preceding the manual response.

5.2.2 Overlaps

Compared to the gaps, there were few overlaps exceeding 1000 ms. Only three overlap transitions in the data set were of that length, one in Group 6, and two in Group 2.

In the conversation between KFV, MS, and ØSR (Group 6), there was one instance of an overlap transition with a duration exceeding 1000 ms. The transition followed a question asked by KFV which was answered by ØSR. KFV starts signing the first part of the question (Figure 5.15 and Figure 5.16) before ØSR starts signing.

Figure 5.15: Snapshots from a long overlap in conversation 6 illustrating KFV's (woman on the left) utterance before the overlap starts (Ferrara and Ringsø, 2021).

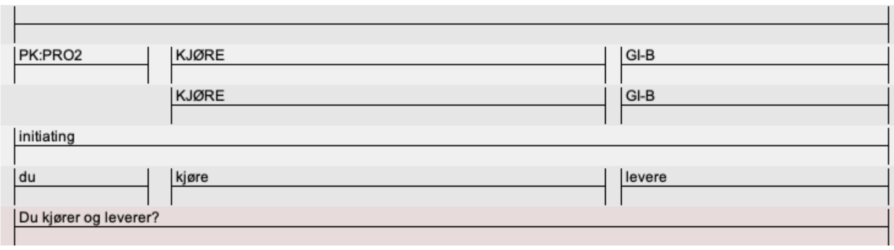


Figure 5.16: Gloss and translations corresponding to the signs depicted in Figure 5.15 (Ferrara and Ringsø, 2021).

KFV signs INDEX (referring to ØSR) DRIVE GIVE 'you drive and deliver?' (Figure 5.15 and Figure 5.16) before ØSR starts replying (Figure 5.17).

Figure 5.17: Snapshots from a long overlap in conversation 6 illustrating the overlapping signing of KFV (woman on the left) and ØSR (man on the right) (Ferrara and Ringsø, 2021).

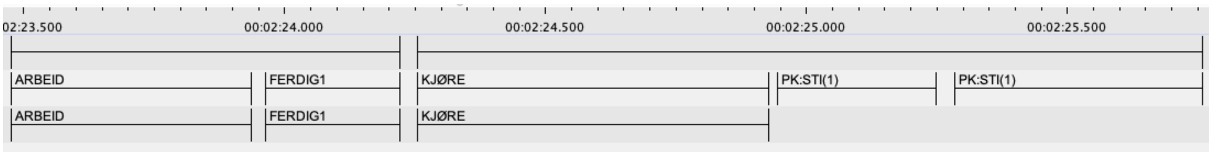


Figure 5.18: Annotations corresponding to the signs depicted in Figure 5.17 (Ferrara and Ringsø, 2021).

While ØSR starts replying, KFV continues signing WORK FINISH DRIVE INDEX (dynamic pointing) 'after work, you drive (location)?', adding to the question (Figure 5.17 and Figure 5.18). It does not seem like ØSR is looking at KFV when he starts to sign the answer to the first part of the question, and thus misses the second part and creates a long overlap.

In the conversation between CJV, and TJ (Group 2), there were two instances of overlap transition durations exceeding 1000 ms. The transitions followed questions asked by CJV and answered by TJ. In the first instance, TJ predicts the end of the question and CJV and TJ sign the last sign together. The second instance is illustrated below (Figure 5.19, Figure 5.20, Figure 5.21, and Figure 5.22).

Figure 5.19: Snapshots of a long overlap from conversation 2 illustrating CJV's (on the left) utterance preceding the overlap (Ferrara and Ringsø, 2021).

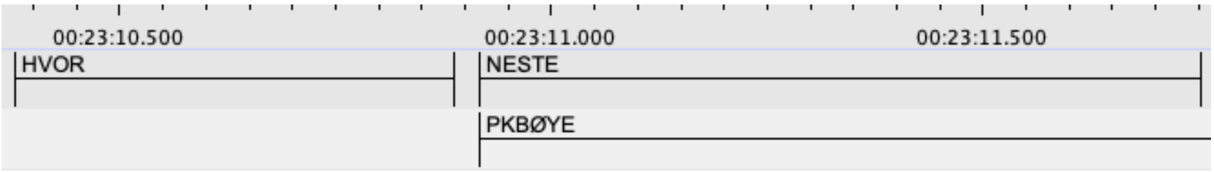


Figure 5.20: Annotations corresponding to the signs depicted in Figure 5.19 (Ferrara and Ringsø, 2021).

As illustrated in Figure 5.19 and 5.20, CJV starts asking a question by signing NEXT WHERE 'where is the next?' (Figure 5.19 and Figure 5.20).

Figure 5.21: Snapshots of a long overlap from conversation 2 illustrating the overlapping signing between CJV (one the left) and TJ (on the right) (Ferrara and Ringsø, 2021).

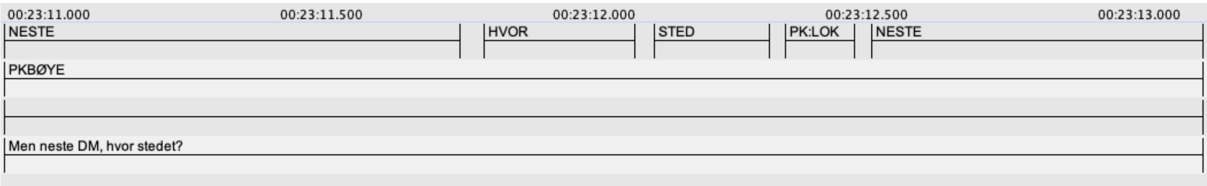


Figure 5.22: Annotations corresponding to the signs in Figure 5.21 (Ferrara and Ringsø, 2021).

While CJV continues, adding to her question, signing WHERE PLACE INDEX (location) NEXT ‘what place is the next?’ (Figure 5.21 and Figure 5.22), TJ is not looking at her while signing his reply to the first part of the question, seemingly missing the second part, and creating a long overlap.

In the overlaps exceeding 1000 ms, two instances were similar, where the questions were asked in two parts, resulting in a response before the second part was uttered. In the remaining instance, the ending of the question seemed to be predicted before the last sign.

6 Discussion

The study presented here is a corpus analysis of NTS with an aim to find out whether turn transition durations in NTS differ from observations of turn-timing in spoken languages, as well as other signed languages to investigate whether the tolerance for overlapping talk and long gaps in NTS is greater or less than in other languages. The focus of the analysis was on question-answer sequences because the context of questions in general requires a response and thus requires the use of the turn-taking system (de Vos et al., 2015, p. 2).

A sign consists of several phases, the preparation phase, the stroke, the hold, and the retraction phase (Kita et al., 1998). Previous research on the topic has indicated that the preparation phase of a sign corresponds to the pre-utterance inbreath before speech (McCleary & de Arantes Leite, 2012), and that the stroke is the phase where the lexical/meaningful information is conveyed. In this study, therefore, stroke-to-stroke turn boundaries were considered when measuring the transition durations of question-answer sequences. The current study was based on a selection of informal conversations with spontaneous interactions from two data sets from the NTS Corpus. Specifically, this study's intent was to find out whether (1) turn transition durations in NTS question-answer sequences are in the range observed for other spoken and signed languages, (2) the average durations are variable between individuals, (3) the average durations are affected by the type of question and (4) the average durations are affected by the age range of the participants. The final data set analyzed in this study included 159 transitions, with 100 gaps, and 59 overlaps.

Previous research on overlap and gaps in conversational turn-taking in spoken languages has found that generally, speakers attempt to minimize overlaps and gaps, and that one speaker contributes to the conversation at a time (Sacks et al., 1974). It has also been found that if there is overlapping talk, there are resolution/repair mechanisms speakers use to resolve the overlap or gap (Schegloff, 2000). Research on turn transition timing in spoken languages also support the claim for minimizing overlap and gaps. A study on transition durations in ten different spoken languages found that the average duration for all languages was 208 ms, and that each of the ten languages measured an average transition duration within a range of 250 ms of the cross-linguistic average (Stivers et al., 2009), approximately the time it takes to produce an English syllable (p. 10588).

Though it has been claimed that signed languages may allow for more overlap than spoken languages, due to the difference in modality (no noise disturbance even during overlap) (Coates and Sutton-Spence, 2001), most research on overlapping talk in signed languages have found that despite the difference in modality, turn-taking and turn-timing follow the same basic principles as spoken languages (de Vos et al., 2015; Girard-Groeber, 2015; Levinson & Torreira, 2015; McCleary & de Arantes Leite, 2012) found that overlap resolution devices are used to terminate or resolve overlap in Brazilian Sign Language (Libras). de Vos et al. (2015) looked at transition durations in question-answer sequences in Sign Language of the Netherlands. They investigated three different turn boundaries but found that only when looking at stroke-to-stroke turn boundaries, the average transition duration fell within the 250 ms range of the cross-linguistic average found in Stivers et al. (2009). When looking at this measure, they found an average transition duration of 307 ms. Girard-Groeber (2015) found that when overlapping talk does occur in Swiss-German Sign Language, it most often happens at transition relevance places which supports the fact that there is a desire for only one speaker at a time.

6.1 Transition durations in NTS

Results from this study indicated that transition durations in NTS falls within the cross-linguistic range found for spoken languages and Sign Language of the Netherlands, measuring an average duration time of 185 ms and a median duration of 138 ms. As mentioned above, this study measured stroke-to-stroke turn boundaries, excluding the preparation phase, the hold, and the retraction phase of the signs. The results may then indicate that, as McCleary and de Arantes Leite (2012) argued, the preparation phase of a sign corresponds to the physical preparation that an inbreath entails in speech and should not be included as part of a turn. However, the conscious choice of including only the stroke in the turns is based on only a few studies and was mainly done to be able to compare transition durations from this study with transition durations found in spoken languages. It was also done because of limited time. Had there not been any limitations of time for conducting this research, it would have been interesting to compare measures of turn boundaries including the preparation phase as well to see whether there are significant differences between signed languages, as this was one of the measures included in de Vos et al. (2015).

The results presented here also support previous findings of one speaker or signer at a time, and that there is a desire for minimizing gaps and overlaps. This can be seen both in the average duration time in milliseconds (185 ms), but also in that, of the 159 transitions investigated in this study, there are in total 100 gaps, but only 59 overlaps. This contradicts Coates and Sutton-Spence (2001) who argued that signed languages have a greater tolerance for overlapping 'talk' than spoken languages in that they found that in British Sign Language, "the collaborative floor" is used for a great deal of the conversations (Coates and Sutton-Spence, 2001). However, this study is currently, to my knowledge, the only one that has looked at overlapping signing in NTS and more research on the topic is needed to challenge the collaborative floor theory.

6.2 Individual differences

Because in the absence of research in NTS, the current study also investigated whether there were individual differences in transition durations between signers. The analysis clearly showed differences in duration depending on who asked the question, and on who answered it. Four participants showed a wider variety of durations when asking questions, but the number of questions asked per participant was unbalanced, and three of the four that showed the greatest variety of durations were also the participants who asked the most questions (38, 37, and 16). However, one other participant who asked a similar number of questions (24), showed less variety, indicating that there was a greater consistency in transition durations when they asked the question. The majority of participants who asked questions, asked less than five questions each. Due to this imbalance, it is difficult to say if there really are individual differences or not as more data per participant is needed to conclude anything.

The same trend could be seen for participants answering questions. Most questions answered were done by a few participants, and half of the participants who answered questions answered less than five each. One participant only answered one question. This imbalance, as mentioned above, makes it difficult to say much about individual differences, but it is still possible to see that there are some.

6.3 The effect of question type and age

Another factor this study looked at was whether the type of question or the age range of the participant influenced the average transition duration. Question types were considered because previous research has found that in English and Swedish spoken conversation, there is a significant difference in response time between question types (Strömbergsson et al., 2013, p.

2587). Three types of questions were included in this study, wh-questions, yes/no-questions, and alternative questions (with 54, 89, and 11 questions respectively). There were an additional five questions in the data set. However, these were questions where the type was uncertain and was therefore excluded from the analysis.

Observationally, alternative questions stood out among the three types in that they had a negative mean transition time. However, the statistical analysis showed that no significant difference was found between the three different types of questions and their effect on transition durations. This does not necessarily mean that question type has no effect on transition duration in NTS. The observational trend of faster response time to alternative questions is in line with previous research, and it might simply mean that there was not enough data to represent each type.

Previous reaction time research has found that age affects response times (Hultsch et al., 2002). We might then expect a somewhat slower response time and longer gaps between question and answer for the upper age ranges. However, no significant difference was found to indicate that in this study. The age range of participants asking and answering the question did not statistically influence transition durations, but this is, however, likely due to the imbalance in number of questions asked and answered by each participant, seeing as the age ranges 50-60 and 60-70 contained less than ten data points each.

6.4 Gaps and overlaps exceeding 1000 ms

As mentioned above, the analysis of this study resulted in an average transition duration time of 185 ms and a median duration of 138 ms. Most cases of overlap and gaps in transitions were very short. However, there were outliers in each category where some gaps and overlaps were very long. 15 transition durations that exceeded 1000 ms were further investigated to see if they shared any characteristics or if there was a reason that could be observed explaining why these transitions were so long. Most of these transitions were gaps (12 transitions), whereas only a few (three transitions) were overlaps.

For almost all the gap transitions, the reason for the long pause was that the person answering the question had to think before replying. In the remaining one gap transition, a non-manual response was given before the manual response, meaning the actual gap between the question and the response was shorter than the measured gap duration. As this study only looked at manual signs, non-manual response was not considered. There might be more cases, for both overlap and gaps, where non-manual response preceded the manual response, which were not investigated in this study. However, previous research on turn-timing, both in spoken and signed languages, also excluded non-manual markers from their studies (de Vos et al., 2015; Stivers et al., 2009), making it a non-issue comparison-wise. Nevertheless, it would be interesting to include non-manual signs in future studies of turn-timing.

There were only three cases of overlaps that exceeded 1000 ms. Overlaps may be due to one or more signer trying to take the next turn, or that the signer asking the question continues beyond a transition relevance place without a repair mechanism being used right away (Levinson and Torreira, 2015). This seems to be the case in two of the examples of long overlaps. It seems that the person answering the question thinks the first signer is done asking the question, but then the first signer continues the question after the second signer has started to reply. This is in line with previous research on overlapping signing where it has been found that overlaps most often happen around transition relevance places (Levinson and Torreira, 2015). In the remaining instance of long overlaps, the person answering the question seemed to predict the last sign of the question and confirms by signing the same sign as the last sign in the question.

6.5 Limitations of the study

One major limitation of this study was the size of the data set. In this study, this came down to limited time and resources. As this was a master's thesis project, there were not enough resources available to go through the entire NTS corpus and a decision had to be made to only include a few data sets that each included a set number of questions per signer. Of the 194 total questions in the files, 159 questions were annotated for the current study. The remaining questions were excluded from the analysis either because they had no response, or because the context of the question was unclear to the researcher who is not a fluent signer of NTS. As mentioned above, the analysis of the effect question type and participant age had on transition duration did not show any significant difference. Because there was a limited amount of data, these results do not mean there definitively was not effect, there was just not enough information in the data set to find out. With a larger data set of questions, the results might have been different.

The results from this study comes from a quantitative analysis and a brief qualitative analysis. The quantitative analysis was done to make the study comparable with Stivers et al. (2009) and de Vos et al. (2015), and the qualitative analysis was done to investigate if there were any important contextual factors involved in long overlaps and gaps. However, by looking only at the context of long gaps and overlaps, the environments surrounding all other gap and overlap transitions were missing. The context surrounding the turn transitions could have said something about why the gaps and overlaps occur, and though we can say that there seems to be a desire for no overlap, no gap, and that long gaps are often due to the participants thinking, we cannot state much further. It is not possible to state whether anything was done in these environments to for instance prevent or end overlap, which was found to be the case in Brazilian Sign Language (McCleary & de Arantes Leite, 2012). It is also impossible to make a generalization about in what specific environments (beyond question type) overlaps and gaps occur and if there is a correlation between duration of transitions and the environments. It would, therefore, be interesting to see what a more detailed qualitative, as well as a quantitative analysis would bring of information in further research of turn-timing in NTS.

7 Conclusion

By looking at turn transition durations in informal conversations, the aim of this thesis was to investigate whether NTS showed a greater tolerance for overlap than what has previously been found for spoken languages by looking at turn-timing in question-answer sequences in NTS. Specifically, to see if the average transition duration falls within the cross-linguistic range observed for other signed and spoken languages and if factors such as the age of the signer or the type of question affected the average duration.

Though there has been an increase in research on signed languages across the world since Stokoe (1960) first argued that American Sign Language was a natural language, there are still areas of sign linguistics, in many languages, that are largely unexplored. One such area is overlaps and gaps in turn-taking. A few studies have been done on British Sign Language (Coates and Sutton-Spence, 2001), Brazilian Sign Language (McCleary & de Arantes Leite, 2012), Swiss-German Sign Language (Girard-Groeber, 2015), Sign Language of the Netherlands (de Vos et al., 2015). These studies focused on different aspects of turn-taking, such as the tolerance for overlapping signing, the environments in which overlap occurs, overlap repair mechanisms, and the timing of turn transitions. Though this research indicates that the turn-taking system in signed languages share the characteristics of turn-taking in spoken languages, conclusions cannot be made unless more research on each aspect of the system in each language is done. NTS is largely unexplored in linguistic research, and especially regarding the topic of turn-taking, and this thesis has been an attempt to fill that gap and add knowledge to the overall topic of turn-taking in signed languages.

By analyzing the NTS Corpus in ELAN and measuring stroke-to-stroke turn boundaries in question-answer sequences, a quantitative analysis was done on transition durations in NTS. A closer look was then taken at instances of gaps or overlaps that exceeded 1000 ms to try to determine if they share any characteristics, or if a reason can be observed as to why they are so long.

The results from this study indicate that transition durations in NTS, with an average transition duration of 185 ms, falls within the same cross-linguistic range found in previous research on the topic, supporting the theory that the system of turn-taking in informal conversation may be a system found in all languages, across modalities. The study did not, however, find a significant difference between question types and their effect on durations, nor between age ranges. Some differences were found in average duration of transitions following the question or preceding the answer of specific individuals. However, the number of questions asked and answered by each participant was not balanced, making it impossible to conclude that there were individual differences. In the instances of long gaps (exceeding 1000 ms), a common contextual characteristic was that the participant responding to the question had to think before signing. There were not enough instances of long overlaps to determine a common context.

To further understand the turn-taking system in NTS, it would be interesting to replicate the study with a much larger data set to see if age and question type influences average transition duration in question-answer sequences, as well as if there are clear individual differences depending on who asks and answers the questions. A more substantial qualitative analysis of all gaps and overlaps would also be useful, to further understand the environments in which overlap, and gaps occur.

Though this study could not state whether there are individual differences or if the age of participants or question type affects the average duration of transitions, the research presented here has added to previous research on transition durations in question-answer sequences in signed languages. With limited previous linguistic research on NTS, this thesis has also given an insight into the turn-taking system in the language, demonstrating that there is a desire for minimal gap and overlap in informal conversation. I hope this thesis brings attention to NTS and inspires further linguistic research.

References

Annotating Norwegian Sign Language corpora. (2022).

Alexandra Kuznetsova, Per B. Brockhoff and Rune H. B. Christensen (2017) lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13), 1–26. doi:10.18637/jss.v082.i13

Armstrong, D. F., & Karchmer, M. A. (2009). William C. Stokoe and the Study of Signed Languages. *Sign Language Studies*, *9*(4), 389–397.

Arnesen, K., Enerstvedt, R. T., Engen, E. A., Engen, T., Høie, G., & Vonen, A. M. (2008). The Linguistic Milieu of Norwegian Children With Hearing Loss. *American Annals of the Deaf*, *153*(1), 65–77.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1). 1–48. https://doi.org/10.18637/jss.v067.i01.

Brugman, H., & Russel, A. (2009). Annotating Multimedia / Multi-modal resources with ELAN. *Proceedings of the Fourth International Conference on Language Resources and Evaluation*.

Coates, J., & Sutton-Spence, R. (2001). *Turn-taking patterns in deaf conversation*. 23.

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Conrad, S. (2002). 4. Corpus Linguistic Approaches for Discourse Analysis. *Annual Review of Applied Linguistics*, *22*, 75–95. http://dx.doi.org/10.1017/S0267190502000041

de Vos, C., Torreira, F., & Levinson, S. C. (2015). Turn-timing in signed conversations: Coordinating stroke-to-stroke turn boundaries. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00268

Erlenkamp, S. (2009). « Gesture verbs ». *CogniTextes. Revue de l'Association Française de Linguistique Cognitive*, *3*, Article 3. https://doi.org/10.4000/cognitextes.250

Erlenkamp, S. (2011a). Grunntegnstilling i norsk tegnspråk. *Norsk lingvistisk tidsskrift : NLT*, *29*(1), 87-.

Erlenkamp, S. (2011b). *Norsk tegnspråk: Helt norsk og veldig annerledes. Skisse av en ny beskrivelses- modell for norsk tegnspråk*. *29*, 12.

Ferrara, L. (2020). Some interactional functions of finger pointing in signed language conversations. *Glossa: A Journal of General Linguistics*, *5*(1). https://doi.org/10.5334/gjgl.993

Ferrara, Lindsay & Bø, Vibeke. 2015 (n.d.). *A pilot corpus of Norwegian Sign Language* Norwegian University of Science and Technology. Unpublished video recordings and annotation files.

Ferrara, L. N., & Halvorsen, R. P. (2017). Depicting and describing meanings with iconic signs in Norwegian Sign Language. *375-399*. https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2475889

Ferrara, L. N., & Nilsson, A.-L. (2017). *Describing spatial layouts as an M2 signed language learner*. https://doi.org/10.1075/sll.20.1.01fer

Ferrara, L., & Ringsø, T. (2019). Spatial Vantage Points in Norwegian Sign Language. *Open Linguistics*, *5*(1), 583–600. https://doi.org/10.1515/opli-2019-0032

Ferrara, Lindsay and Ringsø, Torill, 2021, *Norwegian Sign Language corpus – Depicting Perspective*, Common Language Resources and Technology Infrastructure Norway (CLARINO) Bergen Repository, http://hdl.handle.net/11509/144.

Girard-Groeber, S. (2015). The management of turn transition in signed interaction through the lens of overlaps. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00741

Heldner, M., & Edlund, J. (2010). Pauses, gaps and overlaps in conversations. *Journal of Phonetics*, *38*(4), 555–568. https://doi.org/10.1016/j.wocn.2010.08.002

Hultsch, D. F., MacDonald, S. W. S., & Dixon, R. A. (2002). Variability in Reaction Time Performance of Younger and Older Adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *57*(2), P101–P115. https://doi.org/10.1093/geronb/57.2.P101

Johannessen, J. B., & Erlenkamp, S. (2003). På språkjakt: Problemer og utfordringer i språkvitenskapelig datainnsamling. In *Norbok*. Unipub. https://urn.nb.no/URN:NBN:no-nb_digibok_2016071908009

Johansson, S., & Oksefjell, S. (1998). *Corpora and Cross-linguistic Research: Theory, Method, and Case Studies*. Rodopi.

Johnston, T. (2009). *Creating a Corpus of Auslan within an Australian National Corpus*.

Kita, S., van Gijn, I., & van der Hulst, H. (1998). Movement phases in signs and co-speech gestures, and their transcription by human coders. In I. Wachsmuth & M. Fröhlich (Eds.), *Gesture and Sign Language in Human-Computer Interaction* (Vol. 1371, pp. 23–35). Springer Berlin Heidelberg. https://doi.org/10.1007/BFb0052986

KODA. (2020, February 12). Norges Døveforbund. https://www.doveforbundet.no/tegnsprak/koda

Levinson, S. C., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00731

Lov om språk (språklova), § 7 (2021). https://lovdata.no/dokument/NL/lov/2021-05-21-42

McCleary, L. E., & de Arantes Leite, T. (2012). Turn-taking in Brazilian Sign Language: Evidence from overlap. *Journal of Interactional Research in Communication Disorders*, *4*(1), 123–154. https://doi.org/10.1558/jircd.v4i1.123

Næss, Å. (2021). Arnfinn Muruvik Vonen: "Norsk tegnspråk: en innføring." *Maal og Minne*, *113*(2), Article 2. http://ojs.novus.no/index.php/MOM/article/view/2004

Perniss, P. M., Pfau, R., & Steinbach, M. (2007). *Visible Variation: Comparative Studies on Sign Language Structure* (1. Aufl., Vol. 188). Mouton de Gruyter.

Pfau, R., Steinbach, M., & Woll, B. (Eds.). (2012). *Sign language: An international handbook*. De Gruyter Mouton.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Raanes, E. (2011). Tegnrom og taktilt tegnspråk. *Norsk lingvistisk tidsskrift : NLT*, *29*(1), 54-.

RStudio Team (2022). RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

Ruiter, J.-P. de, Mitterer, Holger., & Enfield, N. J. (2006). Projecting the End of a Speaker's Turn: A Cognitive Cornerstone of Conversation. *Language*, *82*(3), 515–535. https://doi.org/10.1353/lan.2006.0130

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). *A Simplest Systematics for the Organization of Turn-Taking for Conversation*. *50*(4), 40.

Sandler, W., & Lillo-Martin, D. (2001). Natural Sign Languages. In M. Aronoff & J. Ress-Miller (Eds.), *The Handbook of Linguistics* (pp. 533–562). Blackwell Publishers Ltd. https://doi.org/10.1002/9780470756409.ch22

Schegloff, E. A. (2000). Overlapping talk and the organization of turn-taking for conversation. *Language in Society*, *29*(1), 1–63. https://doi.org/10.1017/S0047404500001019

Schroder, B. S. (2011). Imperativ i norsk tegnspråk. *Norsk lingvistisk tidsskrift : NLT*, *29*(1), 136-.

Selting, M. (2000). The construction of units in conversational talk. *Language in Society*, *29*(4), 477–517. https://doi.org/10.1017/S0047404500004012

Selvik, K.-A. (2006). *Spatial paths representing time: A cognitive analysis of temporal expressions in Norwegian sign language: Vol. nr 247*. Faculty of Humanities, University of Oslo Unipub.

Selvik, K.-A. (2011). Tidsuttrykk i rommet: En kognitiv lingvistisk tilnærming til en gruppe tidsuttrykk i norsk tegnspråk. *Norsk lingvistisk tidsskrift : NLT*, *29*(1), 38-.

Sidnell, J., & Stivers, T. (Eds.). (2013). *The handbook of conversation analysis*. Wiley-Blackwell.

Skedsmo, K. (2020). Multiple Other-Initiations of Repair in Norwegian Sign Language. *Open Linguistics*, *6*(1), 532–566. https://doi.org/10.1515/opli-2020-0030
Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., de Ruiter, J. P., Yoon, K.-E., & Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, *106*(26), 10587–10592. https://doi.org/10.1073/pnas.0903616106
Stokoe, W. C. (1960). Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf. *Journal of Deaf Studies and Deaf Education*, *10*(1), 3–37. https://doi.org/10.1093/deafed/eni001
Strömbergsson, S., Hjalmarsson, A., Edlund, J., & House, D. (2013). *Timing responses to questions in dialogue* (p. 2588). https://doi.org/10.21437/Interspeech.2013-581
Vogt-Svendsen, M. (1981). Mouth position & mouth movement in Norwegian Sign Language. *Sign Language Studies*, *33*(33), 363–376.
Vogt-Svendsen, M. (1990). *Interrogative strukturer i norsk tegnspråk: En analyse av nonmanuelle komponenter i 86 spørsmål*. Det historisk-filosofiske fakultet, Den allmennvitenskapelige høgskolen, Universitetet i Trondheim.
Vonen, A. M. (2020). *Norsk tegnspråk: En grunnbok* (1. utgave.). Cappelen Damm akademisk.
Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). *ELAN: a Professional Framework for Multimodality Research*. 4.

Appendix

Below is the table of exported ELAN annotations of overlap and gap transition durations.

Table A.1 The exported transition durations from ELAN.

Annotation	AnnotationBeginTime	AnnotationEndTime	AnnotationDuration
gap; Q: KFV A: MS JN	3115	4093	978
gap; Q: KFV A: ØSR HV	8704	8748	44
overlap; Q: KFV A: ØSR JN	22553	22565	12
gap; Q: KFV A: ØSR JN	24941	25331	390
gap; Q: KFV A: ØSR JN	26320	26433	113
overlap; Q: KFV A: ØSR JN	28124	28274	150
gap; Q: KFV A: MS JN	57972	59082	1110
overlap; Q: KFV A: ØSR JN	66889	66972	83
gap; Q: KFV A: ØSR JN	77041	78017	976
overlap; Q: MS A: ØSR JN	79274	79276	2
gap; Q: KFV A: ØSR HV	91635	92378	743
overlap; Q: ØSR A: KFV JN	96671	96738	67
overlap; Q: KFV A: ØSR JN	119437	119801	364
overlap; Q: KFV A: ØSR JN	144523	145696	1173
gap; Q: MS A: ØSR JN	159171	159583	412
overlap; Q: KFV A: ØSR HV	186279	186383	104
gap; Q: KFV A: ØSR HV	191200	192223	1023
gap; Q: MS A: KFV uncertain	228934	229083	149
overlap; Q: KFV A: ØSR JN	233512	233656	144
overlap; Q: KFV A: ØSR JN	242015	242224	209
gap; Q: KFV A: ØSR HV	244116	244467	351
overlap; Q: KFV A: ØSR JN	251445	252196	751
gap; Q: ØSR A: MS HV	272289	272736	447
overlap; Q: KFV A: MS JN	278794	278854	60
gap; Q: KFV A: MS JN	280187	281055	868
gap; Q: ØSR A: MS HV	307322	307406	84
overlap; Q: KFV A: MS ALT	314439	314454	15
overlap; Q: KFV A: MS JN	346226	346365	139
gap; Q: KFV A: MS HV	347954	348733	779
gap; Q: KFV A: MS JN	382142	382517	375
gap; Q: KFV A: MS JN	391530	392048	518
overlap; Q: KFV A: MS HV	395296	395704	408
gap; Q: KFV A: MS ALT	441389	441662	273
gap; Q: KFV A: MS JN	457940	459015	1075
gap; Q: KFV A: MS JN	460180	461560	1380
gap; Q: KFV A: MS JN	499196	499761	565

gap; Q: KFV A: MS uncertain	500328	501401	1073
overlap; Q: KFV A: MS JN	508217	509106	889
overlap; Q: KFV A: MS ALT	532108	532525	417
gap; Q: KFV A: MS JN	536257	536274	17
gap; Q: KFV A: MS JN	563786	564181	395
gap; Q: KFV A: ØSR HV	580583	580672	89
gap; Q: KFV A: ØSR HV	587395	587458	63
gap; Q: MS A: ØSR JN	693731	694021	290
gap; Q: MS A: ØSR JN	781030	781198	168
overlap; Q: MS A: ØSR JN	783855	784659	804
overlap; Q: MS A: ØSR ALT	790637	791095	458
overlap; Q: MS A: ØSR JN	836317	836857	540
gap; Q: MS A: ØSR JN	875814	876218	404
gap; Q: ØSR A: KFV JN	1318937	1319151	214
gap; Q: MS A: KFV JN	1351044	1351657	613
gap; Q: MS A: KFV JN	1354260	1355169	909
overlap; Q: MS A: ØSR HV	1429842	1430420	578
gap; Q: MS A: ØSR JN	1436478	1436900	422
gap; Q: MS A: ØSR JN	1446112	1446124	12
overlap; Q: MS A: KFV JN	1449323	1449693	370
gap; Q: MS A: ØSR JN	1449694	1449716	22
overlap; Q: IMH A: TH JN	536457	536592	135
overlap; Q: IMH A: TH HV	982427	982615	188
gap; Q: IMH A: TH HV	1622930	1623843	913
gap; Q: ER A: UN2 HV	51945	52043	98
gap; Q: ER A: UN2 HV	53156	53709	553
gap; Q: ER A: UN2 JN	696528	697107	579
overlap; Q: PN A: UN3 HV	97694	97717	23
gap; Q: PN A: UN1 JN	1440855	1441060	205
gap; Q: PN A: UN3 ALT	1602576	1602674	98
gap; Q: PN A: UN3 JN	1635352	1635937	585
overlap; Q: PN A: UN3 JN	1934039	1934201	162
gap; Q: PN A: UN3 JN	2457260	2459024	1764
overlap; Q: BHS A: CJV JN	161087	161366	279
gap; Q: BHS A: CJV JN	344169	344208	39
gap; Q: BHS A: CJV JN	352592	353056	464
gap; Q: CJV A: TJ JN	5975	6347	372
overlap; Q: CJV A: TJ HV	7552	7615	63
gap; Q: TJ A: CJV JN	8211	8767	556
gap; Q: CJV A: TJ HV	13997	14342	345
gap; Q: CJV A: TJ JN	27032	27901	869
gap; Q: CJV A: TJ HV	75663	75801	138
gap; Q: CJV A: TJ JN	82142	82340	198

gap; Q: CJV A: TJ HV	115774	116292	518
overlap; Q: CJV A: TJ JN	126190	126735	545
gap; Q: CJV A: TJ HV	138987	139939	952
overlap; Q: CJV A: TJ HV	173065	173635	570
overlap; Q: CJV A: TJ HV	214718	215412	694
overlap; Q: CJV A: TJ uncertain	372582	372952	370
overlap; Q: TJ A: CJV JN	380721	381424	703
overlap; Q: CJV A: TJ ALT	465947	466728	781
gap; Q: CJV A: TJ JN	722190	723563	1373
overlap; Q: CJV A: TJ HV	760976	761090	114
gap; Q: CJV A: TJ HV	898039	898207	168
overlap; Q: CJV A: TJ HV	972883	973460	577
overlap; Q: CJV A: TJ ALT	995224	996191	967
gap; Q: CJV A: TJ HV	1002329	1002529	200
gap; Q: CJV A: TJ HV	1011185	1012101	916
gap; Q: CJV A: TJ HV	1013706	1014063	357
overlap; Q: CJV A: TJ HV	1092470	1092682	212
overlap; Q: CJV A: TJ JN	1130207	1131675	1468
overlap; Q: CJV A: TJ JN	1151740	1152042	302
gap; Q: CJV A: TJ JN	1177821	1177983	162
overlap; Q: CJV A: TJ JN	1208363	1208832	469
overlap; Q: CJV A: TJ HV	1210367	1211219	852
gap; Q: CJV A: TJ ALT	1227731	1229062	1331
gap; Q: CJV A: TJ ALT	1267028	1267071	43
gap; Q: CJV A: TJ ALT	1289772	1289799	27
gap; Q: CJV A: TJ HV	1296089	1298103	2014
gap; Q: CJV A: TJ HV	1302563	1303087	524
gap; Q: CJV A: TJ JN	1316005	1316025	20
gap; Q: CJV A: TJ HV	1346277	1346320	43
overlap; Q: CJV A: TJ HV	1371954	1372279	325
overlap; Q: CJV A: TJ HV	1392209	1393463	1254
gap; Q: CJV A: TJ HV	1442837	1444650	1813
overlap; Q: CJV A: TJ HV	1509066	1509387	321
gap; Q: EB A: CJV JN	646324	646413	89
gap; Q: EL A: CJV JN	688937	689280	343
overlap; Q: MF A: CJV JN	744963	745013	50
gap; Q: EL A: MF JN	875330	875863	533
gap; Q: MF A: CJV JN	1097145	1097490	345
gap; Q: EL A: CJV JN	1397559	1398445	886
gap; Q: EMN A: UN1 JN	219394	219793	399
overlap; Q: EMN A: UN1 JN	781364	782167	803
overlap; Q: EMN A: UN1 HV	783592	783686	94
gap; Q: OIS A: EB HV	8403	8835	432

gap; Q: OIS A: EB HV	208195	208476	281
gap; Q: OIS A: EB ALT	221782	221825	43
overlap; Q: OIS A: EB uncertain	231856	231931	75
gap; Q: OIS A: EB JN	254732	255627	895
gap; Q: TR A: EB HV	355910	356270	360
overlap; Q: OIS A: EB HV	369467	369479	12
gap; Q: OIS A: EB HV	388417	389149	732
gap; Q: OIS A: EB JN	393812	394019	207
gap; Q: OIS A: EB JN	396356	396870	514
overlap; Q: OIS A: EB JN	401853	401974	121
overlap; Q: OIS A: EB JN	405497	405535	38
gap; Q: OIS A: EB uncertain	408711	409873	1162
overlap; Q: TR A: EB HV	417273	417635	362
gap; Q: TR A: EB JN	427122	427415	293
overlap; Q: TR A: EB JN	437916	438026	110
gap; Q: EB A: TR JN	466915	467010	95
overlap; Q: OIS A: EB HV	567864	567961	97
gap; Q: OIS A: TR HV	619736	619821	85
gap; Q: OIS A: EB HV	708195	708944	749
gap; Q: OIS A: EB JN	713227	713872	645
gap; Q: OIS A: EB JN	718498	718674	176
gap; Q: OIS A: TR HV	744058	744780	722
gap; Q: TR A: EB HV	842771	843047	276
overlap; Q: TR A: EB JN	845641	846359	718
gap; Q: OIS A: TR JN	866437	866673	236
overlap; Q: TR A: EB JN	934777	935192	415
overlap; Q: TR A: EB JN	976955	977465	510
gap; Q: OIS A: EB JN	1203353	1205377	2024
overlap; Q: TR A: EB JN	1366658	1366839	181
gap; Q: OIS A: EB JN	1559237	1559496	259
gap; Q: OIS A: TR HV	1625209	1625621	412
gap; Q: EB A: TR JN	1672116	1672226	110
gap; Q: OIS A: EB JN	1730039	1730726	687
gap; Q: TR A: EB JN	2023434	2023835	401
gap; Q: EB A: OIS JN	2060484	2061302	818
gap; Q: OIS A: EB HV	2072623	2073046	423
gap; Q: TR A: OIS HV	2218034	2218973	939