

Multimodal feedback i social interaktion

PATRIZIA PAGGIO OG COSTANZA NAVARRETTA

I denne artikel analyserer vi sproglig og nonverbal feedback i to multimodale korpora som afspejler forskellige kommunikationssituationer. Det ene består af otte videoptagede såkaldte *'map task'*-samtaler, hvor en af samtalepartnere forklarer den anden ruten på et kort. De to deltagere sidder i forskellige rum uden at kunne se hinanden og kommunikerer igennem høretelefoner. Det andet korpus er en samling af tolv videoptagelser af frie ansigt-til-ansigt-samtaler hvor deltagerne, som ikke kendte hinanden på forhånd, blev optaget mens de stod over for hinanden i et studie og talte sammen. Sproget er dansk i begge korpora. Vi beskriver hvordan sproglig og nonverbal feedback er opmærket og gennemgår distributionen af forskellige typer af feedback-udtryk i de to korpora. Vores analyse viser interessante forskelle i både mængden og typen (mono- kontra multimodal) af feedback, som dels skyldes de forskellige fysiske opsætninger, dels interaktionens natur, som er funktionel i *'map task'*-dialogerne, og som bare er samtale for samtalsens skyld i de frie ansigt-til-ansigt-dialoger.

EMNEORD: multimodal feedback, nonverbale udtryk, interaktionskorpora, hovedbevægelser

1 INTRODUKTION

Det spiller en stor rolle i den sproglige interaktion at man i løbet af en samtale giver hinanden feedback, dvs. at man tilkendegiver at man lytter, at man forstår hvad der bliver sagt, og at man enten er enig eller uenig. Der er selvfølgelig stor variation i hvor meget og hvilken form for feedback man ser afhængigt af individuelle forskelle mellem talere, men også af situationen og interaktionstypen.

I denne artikel analyserer vi den måde hvorpå feedback kommer til udtryk i to forskellige korpora som afspejler forskellige kommuni-

kationssituationer. Det ene består af otte videooptagede samtaler fra DanPASS-korpusset (Grønnum 2006), hvor en af samtalepartnerne forklarer den anden ruten på et kort. Det andet er det danske NOMCO-korpus (Paggio & Navarretta 2016), som består af tolv videooptagelser af frie samtaler. Der er to vigtige forskelle i de to korpora som begge er relevante for feedback, dels at samtalerne har et bestemt formål i DanPASS, hvad det ikke er tilfældet i NOMCO, dels at samtalepartnerne ikke kan se hinanden i det første korpus, hvorimod de står over for hinanden i det andet. Desuden spiller de to samtalepartnere i DanPASS forskellige roller som følge af dialogernes overordnede formål, og denne rollefordeling er, som vi vil se, også udslagsgivende i forhold til brugen af multimodal feedback.

Vi anvender i artiklen en bred definition af feedback forstået på den måde at vi både kigger på talerens og lytterens multimodale feedbackudtryk, men vi begrænser til gengæld vores analyse til korte feedbacktegn, typisk enkelte ord som kan optræde sammen med hovedbevægelser eller andre nonverbale udtryk.

De teoretiske begreber vi anvender i artiklen, defineres i afsnit 2 ud fra den relevante litteratur. Vi fortsætter i afsnit 3 med at beskrive de to korpora og den måde hvorpå de er transskriberede, for i afsnit 4 at forklare hvordan multimodale feedbackudtryk er opmærket, hvor vi også gennemgår nogle udvalgte eksempler. Afsnit 5 består af en kvantitativ analyse af sproglige og multimodale feedbackudtryk i de to korpora og af en sammenligning af de to. Vi slutter af i afsnit 6 med en diskussion af de analyserede mønstre, hvor vi forklarer forskellene ud fra de forskellige kommunikationssituationer de to korpora afspejler.

2 MULTIMODAL FEEDBACK

I en udveksling som er ret typisk for DanPASS-dialogerne, siger guiden: ”hvis du står og kigger mod nord med ryggen til klosteret”, og følgeren nikker en gang mens han siger ”ja”. Det er karakteristisk for disse dialoger at følgeren bruger nik sammen med positive feedbackord for at bekræfte at han eller hun følger med i det som guiden er ved at forklare. I et andet eksempel siger følgeren ”ja, det er helt forkert” mens hun ryster på hovedet. I dette tilfælde refererer hovedrysten til det at følgeren har fulgt en forkert vej gennem kortet, men samtidig

udgør det multimodale udtryk også positiv feedback til det som guiden har sagt, nemlig at følgeren har taget en forkert rute.

I begge tilfælde er der tale om feedback fra lytteren til taleren. Et andet eksempel, denne gang fra en af NOMCO-samtalerne, illustrerer at den person som har ordet, kan bruge et multimodalt udtryk til at fremprovokere en reaktion fra den anden person. De to deltagere er lige begyndt at snakke sammen, og en af dem forklarer at hun er blevet våd på grund af regnvejret og tilføjer ”som du kan se så” efterfulgt af en kort pause og ordet ”ja” samtidig med at hun lægger sit hoved på skrå. Dette multimodale signal opfatter samtalepartneren som en invitation til at komme med noget feedback, og han siger ”okay, jamen” for så at overtage turen (se figur 2 senere i artiklen).

I Allwood et al. (1992) og Allwood (2000) beskrives feedback som den mekanisme deltagerne i en samtale bruger til at etablere og fastholde *kontakt, perception og forståelse* ved hjælp af korte udtryk, med andre ord til at vise at de kan og vil interagere, at de hører og forstår hvad der bliver sagt, samt at de er enige eller uenige. Ifølge forfatterne kan feedback således gives og fremkaldes gennem korte udtryk, som ikke behøver afbryde hovedtråden i diskursen. Det er velkendt at lytteren giver feedbacktegn – disse kaldes også ofte for *backchannelling* (Yngve 1970), men taleren kan også bruge korte signaler til at fremkalde en reaktion hos lytteren. Denne mekanisme regnes for en helt central komponent i den menneskelige kommunikation og for at være afgørende for dens vellykkethed. En indikation af at feedback spiller en central rolle, er at såkaldte feedbackord som *ja, nej, mm* optræder med meget stor hyppighed. Den type feedback disse ord udtrykker, fortolkes forskelligt afhængigt af konteksten og varierer bl.a. i forhold til diskursens polaritet og talehandling (Allwood et al. 1992). For eksempel betyder *ja* ikke det samme efter et spørgsmål som efter et udsagn, og ligeledes kan et *nej* fortolkes forskelligt afhængigt af om det foregående udsagn er positivt eller negativt.

Vigtigheden af feedback i kommunikationen understreges også i mange andre studier, selvom terminologien er noget forskellig fra den vi bruger. Goodwin (1981) anvender fx termen *attentiveness signals* om en række verbale og nonverbale tegn som både bruges af lytteren til at vise opmærksomhed over for hvad taleren siger, og af taleren til at

sikre sig at lytteren følger med i samtalen. Schegloff (1982) betragter brugen af korte sproglige eller nonverbale udtryk, fx nik, som tegn på at kommunikationen har opfyldt sine mål (*achievement and accomplishment in discourse*), mens Stivers (2008) understreger at feedbackudtryk signalerer at taleren og lytteren tilpasser sig hinanden mens de kommunikerer (*alignment*), og at lytteren følger med i hvad taleren udtrykker. Endelig er *ja-* og *nej-*udtryk blevet undersøgt i danske data af fx Steensig & Heinemann (2013), som dog fokuserer deres analyse på tilfælde hvor udtrykkene bliver brugt efter ja/nej-spørgsmål.

Teoretisk kan man forklare feedbacks centrale rolle ud fra en model hvor en samtale anskues som et fælles projekt, og hvor samtalepartnere begge bidrager til den fælles forståelse (*grounding*) og gennem feedback sikrer sig at denne forståelse er til stede (Clark & Schaefer 1989). Den indflydelse som feedback har på kommunikationens vellykkethed, er også blevet undersøgt eksperimentelt. For eksempel fandt Kraut et al. (1982) at jo mere feedback talere fik mens de genfortalte en film, jo bedre fortællinger producerede de, dvs. at deres lyttere bedre kunne forstå og huske hvad de havde hørt. Meget tyder på at feedback har en positiv virkning i vidt forskellige typer social interaktion, fra fjernundervisning (Mottet 2000) til ældrepatientpleje (Caris & Verhallen et al. 1999). En bedre forståelse af hvordan feedback fungerer i den menneskelige kommunikation, er desuden en forudsætning for at kunne skabe troværdige og brugervenlige virtuelle agenter (Morency et al. 2008).

Flere korpusbaserede studier har vist at nonverbale udtryk, især hovedbevægelser og ansigtsudtryk, er en integreret og vigtig del af feedbackmekanismen (McClave 2000, Cerrato 2007, Allwood et al. 2007, Paggio & Navarretta 2013), hvor i hvert fald en af hovedbevægelsernes funktioner er at regulere interaktionen¹, mens ansigtsudtrykkene tilføjer emotionelt indhold. Generelt er nonverbale udtryk med til at skabe *rapport* mellem samtalepartnere, som ifølge Tickle-Degnen & Rosenthal (1990) bygger på indbyrdes opmærksomhed, positive tilkendegivelser og koordination i interaktionen, dvs. evnen til både metaforisk og fysisk at gå i fælles takt.

Poppe et al. (2011) undersøgte receptionen af feedbacksignaler hos

1 For en diskussion af flere funktioner se fx Heylen (2006).

lyttere ved systematisk at variere mængden, typen og timingen af feedbackudtryk hos både menneskelige talere og virtuelle agenter. De påviste at der er en tendens til at både for lidt og for meget feedback opfattes som unaturlig, hvor den ideelle mængde ser ud til at ligge mellem 6 og 12 signaler per minut. Desuden fandt de at lytterne er mere villige til at acceptere nik uanset konteksten, hvorimod de genererede verbale feedbackudtryk fra virtuelle agenter oftere bliver forkastet som unaturlige.

I denne artikel fokuserer vi på hovedbevægelser og den måde de bliver anvendt på af begge partnere til at udveksle feedback i ansigt-til-ansigt-dialoger. Vi inddrager desuden analyser af interaktionen mellem nonverbale udtryk og talesprog. Vi anvender betegnelsen *multimodal feedback* til at referere til feedback som gives eller fremkaldes ved at anvende flere kommunikationsmodaliteter, dels talesprog, dels nonverbale udtryk som hovedbevægelser, skift i kropsholdning og ansigtsudtryk.

3 SOCIAL INTERAKTION I TO DANSKE KORPORA

I denne artikel har vi valgt at sammenligne den måde hvorpå multimodal feedback viser sig i dele af to danske korpora, DanPASS og NOMCO, hvor interaktionen er forskellig på flere måder, fx i forhold til fysisk tilstedeværelse, forhåndskendskab mellem samtalepartnere og hvilken funktion samtalerne skal tjene. En oversigt over de vigtigste ligheder og forskelle kan ses i tabel 1. Oversigten er efterfulgt af en detaljeret beskrivelse af de to korpora hver for sig i afsnittene 3.1 og 3.2.

TABEL 1. KARAKTERISTIKA VED DE TO ANALYSEREDE KORPORA

| | |
|--|---|
| DanPASS-dialoger | NOMCO |
| Dialoger | Dialoger |
| Deltagerne sidder i hver sit lokale | Deltagere står over for hinanden i samme lokale |
| Mikrofon i lyddæmpet lokale | Loftmonterede mikrofoner |
| Kun en af deltagerne (følgeren) filmet | Begge deltagerne filmet af tre forskellige kameraer |
| Deltagerne kan ikke se hinanden | Deltagerne ser hinanden (hele kroppen) |

| | |
|--|---|
| Deltagerne kommunikerer via høretelefoner | Deltagerne kommunikerer ansigt til ansigt |
| Formål med dialogerne: løse en kort-opgave | Formål med dialogerne: lære hinanden at kende |
| Deltagerne spiller forskellige roller | Ingen forudbestemte roller |
| Ingen krav vedr. forhåndskendskab | Deltagerne kender ikke hinanden |

3.1 DanPASS

DanPASS (Danish Phonetically Annotated Spontaneous Speech) (Grønnum 2009) består af både monologer og dialoger; det er udelukkende dialogerne vi beskæftiger os med her. I dialogerne guider en deltager (*guiden*) den anden deltager (*følgeren*) i at finde vej på et kort, og dernæst bytter de to deltagere roller. Der er små forskelle i de kort som guiden og følgeren har fået udleveret, og derfor er opgaven at guide følgeren fra startpunktet til målet ikke helt ligetil. Deltagerne sidder i forskellige lyddæmpede rum, de kan høre hinanden gennem høretelefoner, men de kan ikke se hinanden. Kun en af deltagerne blev videooptaget i hver dialog (figur 1). Det skal bemærkes at de talere som er genkendelige i artiklens figurer, alle har givet eksplicit og informeret samtykke til at billederne bliver brugt i artiklen.

FIGUR 1. FØLGEREN I EN AF DANPASS-VIDOERNE



I DanPASS-projektet (Grønnum 2009) blev korpusset transskriberet og forsynet med fonetiske, prosodiske og andre lingvistiske opmærkninger i Praat-format (Boersma & Weenik 2009). De akustiske signaler er segmenteret i ord, stavelser og prosodiske fraser. Den ortografiske transskription indeholder en markering af betonedede stavelser, tøven og såkaldte tomme og fyldte pauser. Den prosodiske frasekontur er annoteret som tonehøjdereationer mellem de betonedede stavelser i den prosodiske frase. Ordene er også tilføjet en opmærkning af morfosyntaktiske kategorier (*part of speech*).

I vores arbejde med analyse af feedback i DanPASS har vi fokuseret på dialogerne (lyd og videoer) og især på videooptagelserne af deltagerne i rollen som følger, som er de eneste der blev filmet. Den del af korpusset som vi anvender, består af otte videoer og de tilsvarende transskriptioner. Deltagerne er tre mænd og fem kvinder, som hver optræder både som guide og som følger på skift. Videoernes længde varierer meget. Den korteste video er på 4 minutter, mens den længste er på 11 minutter. I alt varer de otte videoer en time. De eksisterende DanPASS-transskriptioner og -opmærkninger blev importeret til værktøjet ANVIL (Kipp 2004), hvorefter vi tilføjede opmærkninger af feedbackfunktionen til feedbackudtryk. Vi vender tilbage til disse udtryk i afsnit 5.1.

I tabel 2 vises antallet af løbende ord og antallet af unikke ord i de otte DanPASS-dialoger. Dernæst vises antallet af løbende ord og antallet af unikke ord produceret af følgeren i de samme dialoger. Endeligt vises gennemsnittet og standardafvigelsen af ord udtalt af hver deltager i rollen som følger per sekund. Betonedede og ubetonede ord betragtes som forskellige, fx er *ja* (betonet) og *ja* (ubetonet) to forskellige løbende ord. Vi angiver frekvensen af ord per sekund da der er stor variation i længden af dialogerne.

TABEL 2. ORD I DANPASS-DIALOGER

| Statistik | Antal |
|-----------------------------|-------|
| Sum af løbende ord | 9744 |
| Sum af løbende ord (guiden) | 6365 |
| Sum af løbende ord (følger) | 3379 |
| Sum af unikke ord | 940 |

| | |
|--|------|
| Sum af unikke ord (følger) | 513 |
| Gennemsnit af løbende ord per taler per sekund (følger) | 0,93 |
| Standardafvigelse af løbende ord per taler/per sekund (følger) | 0,23 |

Tabellen viser at deltagerne taler mere (næsten det dobbelte antal løbende ord) i rollen som guide end i rollen som følgeren. Dette er ikke så overraskende da guiden skal dirigere følgeren. Lignende resultater er blevet fundet i det engelske *HCRC map task*-korpus, som inspirerede til opbygningen af DanPass-korpusset (Thompson et al. 1993). Variationen i videolængden og i antallet af ord i de enkelte dialoger afhænger af hvor hurtigt deltagerne løste kortopgaven.

Vi fandt 466 sproglige feedbackudtryk i følgerdataene. Hovedbevægelserne og ansigtsudtrykkene i videoerne blev også identificeret og annoteret. Hvordan hovedbevægelserne blev opmærket, forklares i afsnit 4.

3.2 NOMCO

Det danske NOMCO-korpus består af tolv videooptagede dialoger med medvirken af i alt tolv forskellige deltagere, seks mænd og seks kvinder, i alderen 21-36. Hver person deltog i to samtaler, en med en mand og en med en kvinde, og hver samtale varede ca. 5 minutter. Hele materialet udgør derfor ca. en times interaktion. Deltagerne stod over for hinanden. De blev filmet med tre kameraer fra tre forskellige vinkler, og deres tale blev optaget via eksterne loftmonterede mikrofoner.

Alle deltagerne gav skriftlig tilladelse til at deres data blev gemt og brugt til forsknings- og undervisningsformål.

Det danske NOMCO-korpus er en del af en samling multimodale korpora som blev udviklet for en række nordiske sprog af det nordiske projekt NOrdic Multimodal COrpora (Paggio et al. 2010). Det var et mål i projektet at de videooptagede samtaler skulle være så naturlige som muligt. Derfor blev det besluttet at deltagerne ikke skulle bære udstyr af nogen art (ikke engang mikrofoner), og at deres kommunikation ikke på forhånd skulle begrænses med roller eller opgaver. Den eneste instruktion de fik, var at prøve at lære hinanden at kende.

Efter optagelserne blev hver deltager bedt om at udfylde et spørgeskema hvor de vurderede den afsluttede samtale i forhold til en række parametre på en skala fra 1 til 5. En opsummering af besvarelsenerne kan ses i tabel 4.

TABEL 4. NOMCO-SAMTALERNES EVALUERING (GENNEMSNIT OG STANDARDAFVIGELSE) (DE ENGELSKKE PARAMETRE BRUGT I SPØRGESKEMAET ER I PARENTES EFTER DE DANSKE OVERSÆTTELSER)

| Parameter | Gennemsnit | Standardafvigelse |
|---------------------------------------|------------|-------------------|
| Fornøjelig (<i>Enjoyable</i>) | 4,42 | 0,72 |
| Intim (<i>Intimate</i>) | 2,71 | 1 |
| Vellidt (<i>Liked</i>) | 4,04 | 0,91 |
| Interessant (<i>Interesting</i>) | 4,17 | 0,76 |
| Indflydelse (<i>Influence</i>) | 3,75 | 0,79 |
| Fri (<i>Free</i>) | 4,13 | 0,74 |
| Ikke påvirket (<i>Not affected</i>) | 3,46 | 1,06 |
| Naturlig (<i>Natural</i>) | 2,33 | 1,05 |
| Behaget (<i>Pleased</i>) | 4,58 | 0,58 |
| Afslappet (<i>Relaxed</i>) | 3,58 | 1,06 |
| Veltilpas (<i>At ease</i>) | 3,83 | 0,82 |
| Tilfreds (<i>Content</i>) | 4,46 | 0,88 |

Deltagerne vurderede samtalerne naturlighed til samlet lidt under 2,5, men de var alligevel ikke særlig påvirket af den noget unaturlige situation (3,46), de følte at de var rimelig afslappede (3,58), fri til at udtrykke sig (4,13) og veltilpasse (3,83). Vi konkluderer på denne basis at det er forsvarligt at udtale sig om ansigt-til-ansigt-dialoginteraktionsmønstre ud fra de videooptagede samtaler.

Efter optagelserne blev samtalerne manuelt transskriberet ved hjælp af Praat og med samme transskriptionskriterier som var blevet anvendt til DanPASS. Det indebærer bl.a. at grænserne mellem ordene er opmærket, og at tryk samt en markering af pauser og feedbackudtryk som *mmm* er en del af transskriptionen. Til forskel fra DanPASS-korpus er der dog i NOMCO ingen prosodisk annotation. Tabel 5 viser optællinger over ordene i korpusset.

TABEL 5. ORD I NOMCO

| Statistik | Antal |
|--|-----------|
| Sum af løbende ord | 18.556.00 |
| Sum af unikke ord | 3.002.00 |
| Løbende ord per taler per sekund (gennemsnit) | 2,51 |
| Løbende ord per taler per sekund (standardafvigelse) | 1 |

Hvis man sammenligner med optællingerne i DanPASS-dialogerne (tabel 2), kan man se at deltagerne i NOMCO-samtalerne producerer næsten tre gange så mange ord per sekund som følgerne i DanPASS-dialogerne. Det skyldes sandsynligvis at selve formålet ved dialogerne i NOMCO er at tale, hvorimod ordstrømmen i DanPASS-dialogerne er underkastet et bestemt formål.

Alle hovedbevægelser, ansigtsudtryk og kropsholdningsskift i samtalerne blev identificeret og annoteret med en række både formmæssige og funktionsmæssige kategorier. Eftersom vores fokus i denne artikel er hovedbevægelserne, vil vi forklare hvordan disse er opmærket i næste afsnit.

4 OPMÆRKNING AF MULTIMODAL FEEDBACK

Hovedbevægelserne og deres feedbackfunktion er blevet opmærket lidt forskelligt i de to korpora. Vi vil her kun beskrive den metode som blev brugt i NOMCO fordi den er baseret på en nyere og mere gennemprøvet model. De kategorier og den procedure som blev anvendt da DanPASS blev opmærket, er ikke væsentligt forskellige omend de er noget simplificerede. Opmærkningen tager udgangspunkt i MUMIN-annotationskemaet (Allwood et al 2007), og består af både formelle og funktionelle kategorier. Følgende hovedbevægelseskategorier er defineret i MUMIN-skemaet og brugt i vores annotation: *Nod* (nik nedad), *Jerk* (nik opad), *Tilt* (hoved på skrå), *HeadBackward* (hoved tilbage), *HeadForward* (hoved frem), *SideTurn* (drej til siden), *Shake* (hovedrysten), *Waggle* (vrik med hovedet), *HeadOther* (andet hoved). Typerne blev defineret med henblik på at kunne fange forskelle som kunne resultere i forskellige funktioner i samtalerne. De abstraherer således fra den fysiske variation der kan være mellem forskellige realiseringer af den samme type (fx kan et nik være lille eller stort). I skemaet skel-

nes der dog imellem enkelte og gentagne bevægelser – igen ud fra den hypotese af disse kan bruges forskelligt i samtaler.

Den funktionelle opmærkning vedrører tre forskellige fænomener, nemlig turtagning og diskursstrukturering foruden feedback. Vi vil her udelukkende fokusere på feedback, hvor der i opmærkningskemaet er defineret tre forskellige træk. Det mest grundlæggende hedder *FeedbackBasic: CPU* (= *Contact, Perception and Understanding*). Trækket tildeles et nonverbalt tegn hvis det bliver brugt til enten at give eller fremkalde feedback. Ud over det er alle feedbackudtryk opmærket med trækket *FeedbackDirection*, som kan være *Give* eller *Elicit* (eller en blanding af de to). Desuden blev et *Agreement* eller *Disagreement*-træk samt en *Emotion* tilføjet hvis relevant. I NOMCO-korpusset blev opmærkningen foretaget af tre forskellige forskningsassistenter, adskillige tests blev lavet for at afprøve den fælles forståelse af annotationskategorierne, og alle tvivlstilfældene blev diskuteret og afgjort i fællesskabet (Paggio & Navarretta 2016). *Intercoder reliability* blevet testet på 10 % af materialet, som to af de involverede kodere annoterede efter et træningsforløb. Enigheden blev målt ved hjælp af Cohen's *kappa* (Cohen 1960), og var 0,63 i gennemsnit for alle trækkene (min. 0,55, maks. 0,69). Resultatet er sammenligneligt med hvad andre har opnået i lignende multimodale korpusannotationer (Cavicchio & Poesio 2009).

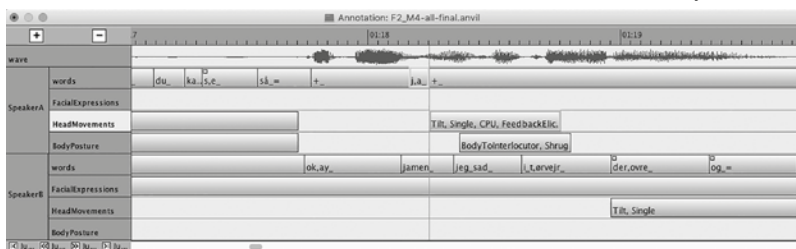
I figur 2 vises to billeder (*frames*) fra en af NOMCO-samtalerne. De illustrerer det eksempel som blev gennemgået tidligere, hvor taleren til venstre fremkalder feedback fra samtalepartneren ved at sige ”ja” samtidig med at hun lægger hovedet let på skrå (svarende til *Tilt* i annotationen). De to gule pile er blevet tilføjet billederne for at indikere hovedbevægelsens retning.

FIGUR 2. EKSEMPEL PÅ MULTIMODALT FEEDBACKUDTRYK I NOMCO

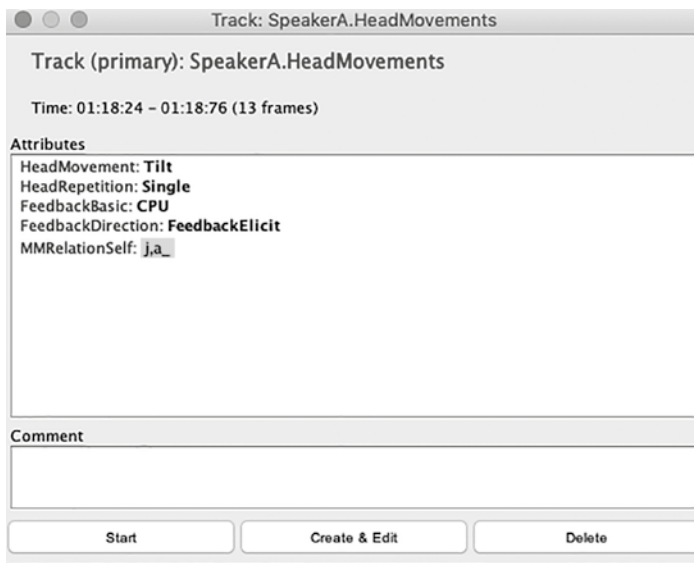


Opmærkning af eksemplet vises i figur 3 og 4, som viser forskellige vinduer fra ANVIL-værktøjet. I figur 3 kan man se de elementer på tidslinjen der svarer til ordene såvel som forskellige typer nonverbale udtryk hos begge talerne. Læg mærke til den markerede hovedbevægelse og det tilhørende *j,a* i ordtransskriptionen for *SpeakerA*. I den anden figur vises de træk som er blevet tilknyttet den markerede hovedbevægelse, dvs. *Tilt*, *Single*, *CPU* og *FeedbackElicit*. Hele annotationen (både ordtransskriptionen og opmærkningen af de nonverbale udtryk) er gemt i XML-format.

FIGUR 3. ORD OG BEVÆGELSESELEMENTER I ANVIL-OPMÆRKNINGSVÆRKTØJET



FIGUR 4. EKSEMPEL PÅ OPMÆRKNING AF HOVEDBEVÆGELSE



5 ANALYSE AF FEEDBACK I DE TO KORPORA

I dette afsnit vil vi analysere hvordan feedback bliver udtrykt i de to korpora. Vi kigger på de sproglige udtryk først og dernæst på hovedbevægelserne, dels alene, dels i kombination med ord.

5.1 Sproglige feedbackudtryk

I tabel 6 vises forekomsten af feedbackudtryk i DanPASS-dialogerne som absolutte tal og som procenter af det samlede ordforråd. Vi viser tallene for begge rollerne sammen og for følgerrollen og guiderollen for sig. Ved at sammenligne tallene i de tre kolonner kan man se at deltagerne i rollen som følger udtrykker verbalt feedback oftere end i rollen som guide.

TABEL 6. FEEDBACKUDTRYK I DANPASS

| Udtryk | Begge roller | | Følgerrolle | | Guiderolle | |
|----------------------|--------------|------|-------------|-------|------------|------|
| | Antal | % | Antal | % | Antal | % |
| Betonet <i>ja</i> | 594 | 6,09 | 377 | 11,15 | 217 | 3,40 |
| Ubetonet <i>ja</i> | 46 | 0,47 | 21 | 0,62 | 25 | 0,39 |
| Betonet <i>jo</i> | 19 | 0,19 | 12 | 0,36 | 6 | 0,09 |
| Ubetonet <i>jo</i> | 17 | 0,17 | 11 | 0,33 | 6 | 0,09 |
| Betonet <i>okay</i> | 112 | 0,15 | 66 | 2 | 46 | 0,72 |
| Ubetonet <i>okay</i> | 1 | 0,01 | 1 | 0,03 | 0 | 0 |
| Betonet <i>nej</i> | 55 | 0,56 | 32 | 0,95 | 23 | 0,36 |
| Ubetonet <i>nej</i> | 18 | 0,18 | 8 | 0,24 | 10 | 0,16 |
| Ubetonet <i>mhm</i> | 48 | 0,49 | 39 | 1,15 | 9 | 0,14 |
| Betonet <i>øh</i> | 3 | 0,03 | 0 | 0 | 3 | 0,05 |
| Alle udtryk | 939 | 9,64 | 579 | 17,14 | 360 | 5,66 |

Tabel 7 viser fordelingen af de forskellige sproglige feedbackudtryk i NOMCO som absolutte tal og procenter af det samlede ordforråd. I alle tilfældene er der tale om udtryk som udgøres af det viste ord eller starter med det, og feedbackudtryk skal forstås således at ordene kan bruges til at udtrykke feedback, dog har vi ikke analyseret hver eneste forekomst.

Hovedparten af udtrykkene udgøres af *ja*-, *jo*- og *nej*-udtryk i deres betonede eller ubetonede form. Disse inkluderer enkelte forekomster, dobbelte udtryk (*ja ja*), og flerordsudtryk som indeholder selve feed-

backordet, og hvor den notation som er brugt i transskriptionen, viser at de enkelte komponenter er blevet udtalt som om der var tale om et ord. Eksempler fra NOMCO er *ja_altså*, *ja_det_er_døjligt*, *jo_sådan*. I disse eksempler indikerer underscoretegnet ‘_’ at der fonetisk er tale om et ord, og ’, ’ viser hvor trykket ligger.

Ud over *ja*-, *jo*- og *nej*-udtrykkene ser vi også en del *okay*. Desuden er *mm/mbm* og *ohm/oh* også betragtet som potentielle feedbackudtryk og talt med. Udtrykket *mbm* bruges i de to korpora som et feedbackord der næsten altid forekommer sammen med et nik. *Oh* udtrykker til gengæld usikkerhed når det anvendes som feedbackudtryk i vores data. Det skal dog bemærkes at *oh* i flere tilfælde bruges når taleren søger efter det korrekte ord snarere end til at udtrykke feedback (Navarretta 2016), og at det derfor har den samme prototypiske funktion som det engelske *uh* (Swerts 1998).

TABEL 7. FEEDBACKUDTRYK I NOMCO

| Udtryk | Antal | % |
|----------------------|-------|------|
| Ubetonet <i>ja</i> | 197 | 1,06 |
| Betonet <i>ja</i> | 581 | 3,13 |
| Ubetonet <i>jo</i> | 35 | 0,19 |
| Betonet <i>jo</i> | 34 | 0,18 |
| Ubetonet <i>okay</i> | 8 | 0,04 |
| Betonet <i>okay</i> | 168 | 0,91 |
| Ubetonet <i>nej</i> | 22 | 0,12 |
| Betonet <i>nej</i> | 106 | 0,57 |
| <i>mm/mbm</i> | 111 | 0,60 |
| <i>ohm/oh</i> | 414 | 2,23 |
| Alle udtryk | 1676 | 9,03 |

Hvis vi sammenligner med DanPASS, kan vi se at den samlede proportion af sproglige feedbackudtryk ligner den for DanPASS-deltagerne i begge roller. Fordelingen af typerne er dog lidt forskellig, hvor betonet *ja* forekommer oftere i DanPASS. Forklaringen er nok at det er meget vigtigt for deltagerne i DanPASS hele tiden at afstemme deres forståelse af opgaven.

5.2 Hovedbevægelser og multimodal feedback

I DanPASS-følger-videoerne blev der kun fundet 139 hovedbevægelser, og 108 (78%) af dem er blevet opmærket med feedbackfunktionen. Alle feedbackhovedbevægelser er multimodale i disse videoer, dvs. de forekommer altid sammen med et sprogligt feedbackudtryk. Dette er ikke overraskende givet at de to samtalepartnere ikke kan se hinanden.

NOMCO-korpusset indeholder i alt 3.117 hovedbevægelser, dvs. i gennemsnit 129,88 hovedbevægelser per taler i hver samtale (SD = 34,59), og 38% af dem (= 1173) er blevet opmærket med feedbackfunktionen. Både her og i andre tilsvarende tabeller har vi bevaret de engelske betegnelser som er brugt i annotationskemaet.

TABEL 8. HOVEDBEVÆGELSER BRUGT TIL FEEDBACK I BEGGE KORPORA (OPMÆRKNINGSKATEGORIERNE ER VIST I PARENTES EFTER DE DANSKE OVERSÆTTELSER)

| Hovedbevægelse | DanPASS | | NOMCO | |
|---------------------------------------|---------|-----|-------|-----|
| | Antal | % | Antal | % |
| Nik opad (<i>Nod</i>) | 68 | 63 | 508 | 43 |
| Nik nedad (<i>JerK</i>) | 13 | 12 | 129 | 11 |
| Hoved på skrå (<i>Tilt</i>) | 11 | 10 | 124 | 11 |
| Hoved tilbage (<i>HeadBackward</i>) | 3 | 3 | 116 | 10 |
| Hoved frem (<i>HeadForward</i>) | 0 | 0 | 85 | 7 |
| Drej til siden (<i>SideTurn</i>) | 2 | 2 | 77 | 7 |
| Hovedrysten (<i>Shake</i>) | 10 | 9 | 70 | 6 |
| Vrik med hovedet (<i>Waggle</i>) | 1 | 1 | 14 | 1 |
| Andet hoved (<i>HeadOther</i>) | 0 | 0 | 50 | 4 |
| Sum | 108 | 100 | 1173 | 100 |

En sammenligning af distributionerne i de to korpora viser et mønster der ligner det vi så for feedbackordenes vedkommende. Vi ser nemlig at nik optræder relativt mere i DanPASS end i NOMCO, ligesom betonet *ja* gør. I begge tilfældene er der tale om et større behov i DanPASS-dialogerne for at udtrykke at man forstår den anden persons beskeder. Hovedrysten er relativt sjælden i begge korpora. I NOMCO er det muligvis på grund af situationen, hvor samtalepartnerne helst ikke vil modsige hinanden. I DanPASS viser det sig at opgaven ikke er alt for

svær at løse. Interessant nok er der faktisk enkelte eksempler i begge korpora hvor hovedrysten bliver brugt til at signalere enighed omend i en negativ kontekst.

Vi har også kigget på hvor ofte hovedbevægelserne forekommer i en sekvens af hurtigt gentagne bevægelser eller som enkeltstående udtryk. I tabel 9 vises denne distinktion kun hvad nik og rysten angår (de andre bevægelser er stort set altid opmærket med *Single*). I DanPASS er der relativt få gentagne bevægelser. I NOMCO er der derimod en stor sandsynlighed for at både nik og hovedrysten optræder i sekvenser, hvorimod det typisk skandinaviske opadgående nik næsten altid ses som en enkelt isoleret bevægelse. En forklaring kan være at det er forholdsvis krævende at producere sekvenser af opadgående nik. Men der kunne også være funktionelle årsager som man bør kigge nærmere på. Fordelingen i NOMCO er signifikant afhængig af skelen mellem enkelt (*single*) og gentaget (*repeated*) bevægelse (chi-i-anden: 199,26, df = 2, p-værdi < 2,2e-16). En forklaring på hvorfor vi ser gentagne bevægelser i NOMCO-dialogerne, er at samtalepartnerne ofte viser deres nonverbale feedback igennem dialogpartnerens længere bidrag snarere end som i DanPASS i form af hurtig feedback til en forespørgsel eller anmodning.

TABEL 9. ENKELTE OG GENTAGNE HOVEDBEVÆGELSER I BEGGE KORPORA, ABSOLUT-TAL (PROCENTER) (OPMÆRKNINGSKATEGORIERNE ER VIST I PARENTES EFTER DE DANSKE OVERSÆTTELSER)

| Hovedbevægelse | DanPASS | | | NOMCO | | |
|------------------------------|----------|----------|----------|----------|----------|-----------|
| | Enkelt | Gentaget | Sum | Enkelt | Gentaget | Sum |
| Nik nedad (<i>Nod</i>) | 46 (68) | 22 (32) | 68 (100) | 160 (31) | 348 (69) | 508 (100) |
| Nik nedad (<i>Jerke</i>) | 13 (100) | 0 (0) | 13 (100) | 127 (98) | 2 (2) | 129 (100) |
| Hovedrysten (<i>Shake</i>) | 4 (40) | 6 (60) | 10 (100) | 17 (24) | 53 (76) | 70 (100) |

I DanPASS-dialogerne er alle feedbackhovedbevægelserne brugt sammen med sproglige udtryk. Dette er ikke tilfældet i NOMCO, og vi har derfor for hver type kigget på om bevægelserne optræder sammen med feedbackord eller -fraser, med andre sproglige udtryk eller alene. Fordelingen vises i tabel 10.

TABEL 10. MONO- OG MULTIMODAL FEEDBACK I NOMCO: ABSOLUTTE TAL (PROCENTER) (OPMÆRKNINGSKATEGORIERNE ER VIST I PARENTES EFTER DE DANSKE OVERSÆTTelser)

| Hovedbevægelse | Feedbackord | Andre ord | Ingen ord | Sum |
|---------------------------------------|-------------|-----------|-----------|------------|
| Gentaget nik (<i>Repeated Nod</i>) | 175 (50) | 36 (10) | 137 (40) | 348 (100) |
| Enkelt nik (<i>Single Nod</i>) | 84 (53) | 37 (23) | 39 (24) | 160 (100) |
| Nik opad (<i>Jerke</i>) | 92 (71) | 22 (17) | 15 (12) | 129 (100) |
| Hoved på skrå (<i>Tilt</i>) | 38 (31) | 52 (42) | 34 (27) | 124 (100) |
| Hoved tilbage (<i>HeadBackward</i>) | 54 (46) | 45 (39) | 17 (15) | 116 (100) |
| Hoved frem (<i>HeadForward</i>) | 25 (29) | 50 (59) | 10 (12) | 85 (100) |
| Drej til siden (<i>SideTurn</i>) | 33 (43) | 34 (44) | 10 (13) | 77 (100) |
| Hovedrysten (<i>Shake</i>) | 31 (44) | 29 (42) | 10 (14) | 70 (100) |
| Vrik med hovedet (<i>Waggle</i>) | 5 (36) | 7 (50) | 2 (14) | 14 (100) |
| Andet hoved (<i>HeadOther</i>) | 14 (28) | 26 (52) | 10 (2) | 50 (100) |
| Sum | 551 (47) | 338 (29) | 284 (24) | 1173 (100) |

Vi har i tabellen valgt at vise 'Enkelt nik' og 'Gentaget nik' separat, hvorimod distinktionen er droppet hvad angår 'Nik opad' og 'Hovedrysten' for ikke at ende med for små tal i de enkelte celler. Fordelingen er også i dette tilfælde signifikant (χ^2 -anden = 203,47, $df = 16$, p -værdi < 0,001)². De største effekter skyldes at 'Gentaget nik' optræder som monomodal feedbacksignal en del hyppigere end andre typer hovedbevægelse, samt at Hoved frem har en stærkere tendens end andre typer til at optræde sammen med sproglige udtryk, dog ikke deciderede feedbackord. Når det optræder med ikke feedbackudtryk, bruges 'Hoved frem' i 70 % af tilfældene til at fremkalde feedback hos den anden taler, hvorimod 'Nik' generelt bruges til at give feedback 90 % af tiden (75 % af tiden når bevægelsen følges af ikke feedbackord).

5.3 Multimodale feedbackudtryk i de to korpora

De sidste optællinger vi vil gennemgå, handler om forekomsten af multimodale feedbackudtryk (hovedbevægelser og tale) i de to korpora og vises i tabel 11. Man skal bemærke at 'taler' i NOMCO betyder alle

2 Tallene for 'Vrik med hovedet' blev talt sammen med dem for 'Andet hoved' inden 'chi-i-anden'-testen for at undgå for små observationer (<5).

deltagerne i alle samtalerne, hvorimod den i DanPASS refererer til deltagerne når de optræder i følgerrollen.

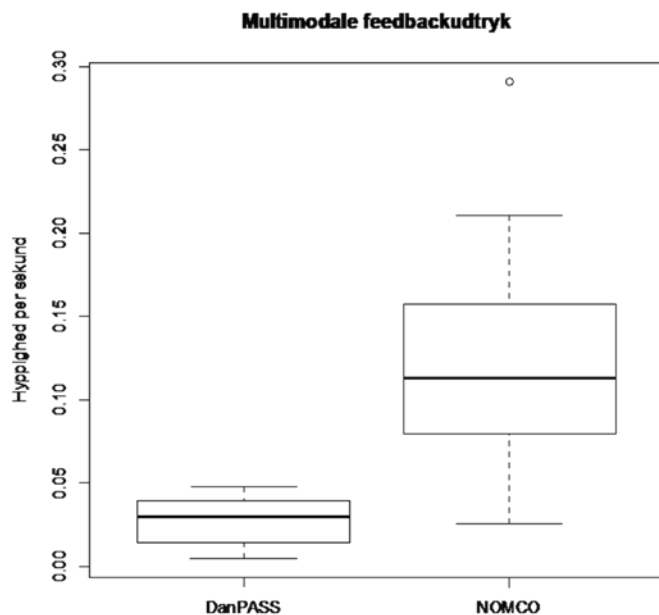
TABEL 11. MULTIMODAL FEEDBACK I DE TO KORPORA

| Statistik | DanPASS (følgerdata) | NOMCO |
|--|-------------------------|-------|
| Sum af multimodale feedbackudtryk | 108 | 889 |
| Multimodal feedback/sproglig feedback | 0,16 | 0,53 |
| Multimodale feedbackudtryk per sekund per taler (gnmsn.) | 0,03 | 0,12 |
| Multimodale feedbackudtryk per sekund per taler (sd) | 0,015 | 0,059 |

For det første kan man se at andelen af multimodal feedback i forhold til sproglige feedbackudtryk er en del større i NOMCO (svarende til 53 %) end i DanPASS-følger-replikkerne (svarende til 16 %). For det andet ser vi også at mængden af multimodal feedback per sekund igen er større i NOMCO, hvor der er 4 gange så mange multimodale feedbackudtryk der involverer hovedbevægelser sammen med tale. Disse forskelle er selvfølgelig ikke overraskende hvis man tager i betragtning at DanPASS-deltagerne ikke kan se hinanden, hvilket de kan i NOMCO. Gennemsnittet på 0,12 multimodale udtryk per sekund (7,32 i minut) som vi ser i NOMCO, falder i øvrigt inden for den rækkevidde på 6 til 12 feedbackudtryk per minut som Poppe et al. (2001) betragter som mest naturlig i en normal samtale. Den lavere frekvens i DanPASS viser at man ikke kan opstille generelle standarder for hvor meget feedback man kan forvente i forskellige typer sproglig interaktion.

I figur 5 vises distributionen af multimodale feedbackudtryk per sekund per taler i de to korpora. Ud over det at der i gennemsnit er færre i DanPASS end i NOMCO, viser plottet at der i NOMCO er større variation imellem talerne, hvor en person endda producerer forholdsmæssigt mere multimodal feedback end alle de andre deltagere. En forklaring på denne variation kan findes i at samtalerne i NOMCO, som vi har set, er meget friere end dem i DanPASS, og det betyder muligvis at forskelle der skyldes individuelle træk eller forskellige grader af afstemning mellem samtalepartnerne, nemmere kommer til udtryk. Forskellen mellem de to distributioner er statistisk signifikant. Signifikansen blev testet dels ved en 'Kruskal-Wallis rank sum'-test på hele

FIGUR 5. DISTRIBUTION AF MULTIMODALE FEEDBACKSIGNALER I DE TO KORPORA



datasættet (chi-i-anden = 15,341, $df = 1$, p -værdi $< 0,001$)³, dels ved en enkeltvejs ANOVA (producerede feedback x korpus) på et reduceret datasæt efter at dataene fra den afvigende taler var blevet fjernet ($F(1, 29) = 23,77$, p -værdi $< 0,001$).

6 DISKUSSION OG KONKLUSION

Analysen af ordfrekvensen i DanPASS og NOMCO viser at der er forskelle i de to korpora som skyldes de forskellige kommunikationssituationer. Først og fremmest siges der flere ord per sekund i NOMCO end i DanPASS. Deltagerne i DanPASS skal udføre en opgave og bruger tid til at se på de to kort som de skal arbejde med. Derimod skal deltagerne i NOMCO bare tale sammen, og de skiftes frit til at spørge om den anden deltagers liv og til at fortælle om sig selv. DanPASS-deltagerne taler i øvrigt mere i guiderollen end i følgerrollen, da de i den første rolle skal dirigere den anden deltager uden at kunne se hinanden

3 Denne test blev valgt på grund af den noget unormale distribution.

og hinandens kort. Denne forskel er også blevet fundet i lignende korpora på andre sprog, bl.a. engelsk (Thompson et al. 1993).

Analysen af feedbackudtryk i de to korpora viser både ligheder og forskelle. I begge korpora udtrykkes feedback hovedsagelig med *ja*- og *nej*-udtryk og med hovedbevægelser. De verbale udtryk og hovedbevægelserne er stort set de samme selvom dialogtypen er forskellig, men deres relative frekvens varierer lidt. Således optræder *nik* og *betonet ja* relativt tiere i DanPASS end i NOMCO som følge af at deltagerne har brug for at tjekke deres fælles forståelse af den opgave de skal løse. I begge korpora optræder hovedrysten mere sjældent end *nik* da deltagerne i begge situationer er mere villige til at bekræfte end modsige hinanden.

Sammenligningen af de to korpora viser at antallet af verbale feedbackudtryk målt i forhold til alle ord i de to korpora er relativt det samme, men analysen af DanPASS-korpusset viser at verbale feedbackudtryk er meget mere frekvente når deltagerne optræder i følgerollen end når de udfylder guiderollen. Dette er ikke overraskende da følgeren er nødt til at give løbende feedback på de instrukser som han eller hun får af guiden.

Den største forskel mellem de to korpora vedrører brugen af hovedbevægelser som feedbacksignal. I NOMCO bruger deltagerne til tider hovedbevægelser uden tale til at udtrykke feedback. Dette er aldrig tilfældet i DanPASS, hvor deltagerne ikke kan se hinanden, og derfor optræder feedbackhovedbevægelserne altid sammen med de verbale feedbackudtryk. Det er dog interessant at DanPASS-deltagerne bruger multimodale feedbackudtryk med en vis frekvens selvom de er alene i et rum og ikke kan ses af den anden deltager. Dette bekræfter at mennesker bruger gestik per refleks når de kommunikerer, og derfor også producerer nonverbale udtryk når de ikke ser deres samtalepartner. Multimodal feedback optræder dog som forventeligt i større grad i NOMCO, dels fordi deltagerne kan se hinanden (se Mol et al. 2011 for relevante eksperimentelle data), dels fordi den friere samtalsituation, som har selve samtalen som mål, kræver at deltagerne hele tiden viser hinanden deres vilje til at bringe samtalen videre. Desuden producerer NOMCO-talerne flere gentagne feedbacksignaler som optræder sammen med dialogpartnerens længere sproglige bidrag.

Det er vigtigt at understrege at vores analyse og konklusioner er baseret på opmærkningskategorier og -procedurer som er teoretisk funderede og afprøvede i praksis på adskillige multimodale samlinger. I forhold til denne artikel har forskellige kodere (2-3) uafhængigt af hinanden genkendt adskillige dialogakter i relation til *ja*- og *nej*-udtryk i DanPASS og flere typer af feedbackrelaterede hovedbevægelser i begge korpora.

Et interessant aspekt vi ikke har undersøgt endnu, vedrører timingen imellem de sproglige og de nonverbale udtryk som sammen skaber de multimodale signaler. En sammenligning af enkelte multimodale udtryk i de to korpora viser at tidsforskydningen mellem hovedbevægelserne og talen, hvor hovedbevægelsen typisk starter først, er større i DanPASS end i NOMCO. Det er som om deltagerne, mens de bevæger hovedet, opdager at deres kropssignal ikke kan blive set af den anden deltager, og derfor skynder de sig også at udtrykke feedback verbalt. I fremtiden vil vi analysere den tidsmæssige relation mellem det verbale og det bevægelsesmæssige feedbackudtryk i alle data for at kunne bekræfte denne observation.

Den generelle konklusion må foreløbig være at multimodal feedback forekommer i begge de korpora vi har analyseret, om end den er en del hyppigere i samtaler hvor deltagerne kan se hinanden, og hvor de taler frit. De individuelle forskelle i mængden og typen af feedback som er til stede i de to korpora, skyldes individuel variation i de frie ansigt-til-ansigt-dialoger, hvorimod den rolle som en deltager har i 'map task'-dialogerne påvirker produktionen af multimodal feedback systematisk.

Det er vores hensigt med denne artikel at bidrage med en empirisk funderet undersøgelse som understøtter vores generelle viden om hvordan hovedbevægelser bruges i samtaler. Undersøgelsen adskiller sig fra og komplementerer kvalitative tilgange til emnet ved at forholde sig systematisk og kvantificerbart til det eksisterende datamateriale. Den er således relevant for sprog- og kommunikationsforskere som beskæftiger sig med gestik og multimodal kommunikation. Men undersøgelsens resultater vil også kunne anvendes af sprogteknologer til at udvikle mere naturlige virtuelle agenter som både kan tolke brugernes gestik og selv generere naturlige nonverbale tegn.

Patrizia Paggio
Center for Sprogteknologi
Institut for Nordiske Studier og Sprogvidenskab, Københavns Universitet
Institute for Linguistics and Language Technology, University of Malta
paggio@hum.ku.dk

Costanza Navarretta
Center for Sprogteknologi
Institut for Nordiske Studier og Sprogvidenskab, Københavns Universitet
costanza@hum.ku.dk

LITTERATURLISTE

- Allwood, J., J. Nivre & E. Ahlsén. 1992. On the semantics and pragmatics of linguistic feedback. *Journal of Semantics* 9(1). 1–26. DOI: 10.1093/jos/9.1.1.
- Allwood, J. 2000. Structure of dialog. M. Taylor, D. Bouwhuis, & F. Neel (red.), *The structure of multimodal dialogue II*, 3–24. Amsterdam: John Benjamins Publishing Company.
- Allwood, J., L. Cerrato, K. Jokinen, C. Navarretta & P. Paggio. 2007. The MUMIN coding scheme for the annotation of feedback, turn management and sequencing. *Multimodal corpora for modelling human multimodal behaviour, Special issue of the International Journal of Language Resources and Evaluation* 41. 273–287. DOI: 10.1007/s10579-007-9061-5.
- Allwood, J., S. Kopp, K. Grammer, E. Ahlsén, E. Oberzaucher & M. Koppensteiner. 2007. The analysis of embodied communicative feedback in multimodal corpora: A prerequisite for behavior simulation. *Language Resources and Evaluation* 4(3–4). 255–272. DOI: 10.1007/s10579-007-9056-2.
- Caris-Verhallen, W.M., A. Kerkstra & J.M. Bensing. 1999. Non-verbal behaviour in nurse–elderly patient communication. *Journal of Advanced Nursing* 29(4), 808–818. DOI: 10.1046/j.1365-2648.1999.00965.x.
- Cavicchio, F. & M. Poesio. 2009. Multimodal corpora annotation: Validation methods to assess coding scheme reliability. M. Kipp, J.C. Martin, P. Paggio & D. Heyen (red.), *Multimodal corpora* (Lecture notes in computer science 5509), 109–121. Berlin: Springer. DOI: 10.1007/978-3-642-04793-0.

- Cerrato, L. 2007. Investigating communicative feedback phenomena across languages and modalities. Ph.d.-afhandling, School of Speech and Music Communication, Stockholm, KTH.
- Clark, H.H. & E.F. Schaefer. 1989. Contributing to Discourse. *Cognitive Science* 13. 259–294. DOI: 10.1207/s15516709cog1302_7.
- Cohen J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20(1). 37–46. DOI: 10.1177/001316446002000104.
- Goodwin, C. 1981. *Conversational organization: Interaction between speakers and hearers*. London: Academic Press. DOI: 10.1525/ae.1983.10.2.02a00430.
- Grønnum, N. 2006. DanPASS – a Danish phonetically annotated spontaneous speech corpus. N. Calzolari, K. Choukri, A. Gangemi, B. Maegaard, J. Mariani, J. Odiijk, D. Tapias (red.), *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC'06)*, 1578–1583. Luxemburg: European Language Resources Association (ELRA).
- Grønnum, N. 2009. A Danish phonetically annotated spontaneous speech corpus (DanPASS). *Speech Communication* 51(7). 594–603. DOI: 10.1016/j.specom.2008.11.002.
- Heylen, D. 2006. Head gestures, gaze and the principles of conversational structure *International Journal of Humanoid Robotics* 3(3). 241–267. DOI: 10.1142/S0219843606000746.
- Kraut, R.E., S.H. Lewis & L.W. Swezey. 1982. Listener responsiveness and the coordination of conversation. *Journal of Personality and Social Psychology* 43(4). 718–731. DOI: 10.1037/0022-3514.43.4.718.
- McClave, E.Z. 2000. Linguistic functions of head movements in the context of speech. *Journal of Pragmatics* 32(7). 855–878. DOI: 10.1016/S0378-2166(99)00079-X.
- Mol, L., E. Krahmer, A. Maes & M. Swerts. 2011. Seeing and being seen: The effects on gesture production. *Journal of Computer-Mediated Communication* 17(1). 77–100. DOI: 10.1111/j.1083-6101.2011.01558.x.
- Morency, L-P., I.A. de Kok & J. Gratch. 2008. Predicting listener backchannels: A probabilistic multimodal approach. H. Prendinger, J. Lester & M. Ishizuka (red.), *Proceedings of the Eight International Conference on Intelligent Virtual Agents 2008*, 176–190. DOI: 10.1007/978-3-540-85483-8_18.
- Mottet, T.P. 2000. Interactive television instructors' perceptions of students' non-verbal responsiveness and their influence on distance teaching. *Communication Education* 49(2). 146–164. DOI: 10.1080/03634520009379202.

- Navarretta, C. 2016. The functions of fillers, filled pauses and co-occurring gestures in Danish dyadic conversations. *Proceedings of the 3rd European Symposium on Multimodal Communication*, 55-61. Linköping: Linköping University Electronic Press.
- Navarretta, C. & P. Paggio. 2010. Classification of feedback expressions in multimodal data. *ACL '10: Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, 318–324. Stroudsburg: Association for Computational Linguistics.
- Paggio, P., J. Allwood, E. Ahlsén, K. Jokinen & C. Navarretta. 2010. The NOMCO multimodal Nordic resource – Goals and characteristics. Calzolari et al. (red.), *Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC'10)*, 2968–2974. Luxembourg: European Language Resources Association (ELRA).
- Paggio, P. & C. Navarretta. 2013. Head movements, facial expressions and feedback in conversations – Empirical evidence from Danish multimodal data. *Journal on Multimodal User Interfaces – Special Issue on Multimodal Corpora* 7(1–2). 29–37. DOI: 10.1007/s12193-012-0105-9.
- Paggio, P. & C. Navarretta. 2016. The Danish NOMCO corpus: Multimodal interaction in first acquaintance conversations. *Language Resources and Evaluation* 51(2). 463–494. DOI: 10.1007/s10579-016-9371-6.
- Poppe R., K.P. Truong & D. Heylen. 2011. Backchannels: Quantity, type and timing matters. H.H. Vilhjálmsón, S. Kopp, S. Marsella & K.R. Thórisson (red.), *Intelligent Virtual Agents. IVA 2011* (Lecture Notes in Computer Science 6895), 228-239. Berlin: Springer. DOI: 10.1007/978-3-642-23974-8_25.
- Schegloff, E.A. 1982. Discourse as an interactional achievement: Some uses of ‘uh huh’ and other things that come between sentences. D. Tannen (red.), *Analyzing discourse: Text and talk*, 71–93. Georgetown: Georgetown University Press.
- Steensig, J. & T. Heinemann. 2013. When ‘yes’ is not enough – as an answer to a yes/no question. B.S. Reed & G. Raymond (red.), *Units of Talk – Units of Action*, 207–242. Amsterdam: John Benjamins Publishing Company. DOI: 10.1075/slsi.25.07ste.
- Swerts, M. 1998. Filled pauses as markers of discourse structure. *Journal of Pragmatics* 30. 485–496. DOI: 10.1016/S0378-2166(98)00014-9.
- Stivers, T. 2008. Stance, alignment, and affiliation during storytelling: When nodding is a token of affiliation. *Research on Language & Social Interaction* 41(1). 31-57. DOI: 10.1080/08351810701691123.

- Tickle-Degnen, L. & R. Rosenthal. 1990. The nature of rapport and its nonverbal correlates. *Psychological Inquiry* 1(4). 285–293. DOI: 10.1207/s15327965pli0104_1.
- Thompson, H.S., A. Anderson, E. Gurman Bard, G. Doherty-Sneddon, A. Newlands & C. Sotillo. 1993. The HCRC map task corpus: natural dialogue for speech recognition. *Proceedings of the 1993 Workshop on Human Language Technology (HLT'93)*, 25–30. Princeton: Association for Computational Linguistics. DOI: 10.3115/1075671.1075677.
- Yngve, V. 1970. On getting a word in edgewise. *Papers from the Sixth Regional Meeting of the Chicago Linguistic Society*, 567-578. Chicago: University of Chicago.