



SIGNSPEAK

**Scientific understanding and vision-based technological development
for continuous sign language recognition and translation**

Grant Agreement Number 231424

**Small or medium-scale focused research project (STREP) FP7-ICT-2007-3.
Cognitive Systems, Interaction, Robotics**

Project start date: 1 April 2009

Project duration: 36 months

Deliverable D.1.8: Additional video recordings for NGT

Dissemination Level: **Public**

Authors: Onno Crasborn & Ellen Ormel
Centre for Language Studies
Radboud University Nijmegen

Version 1, 14 December 2010

Introduction

A small set of new video recordings for NGT was created with the purpose of raising the amount of tokens per type for a number of signs that in the currently annotated set did not reach the minimum of 10 realisations per sign. As opposed to Deliverable 1.6, where sentences were literally repeated, this time new sentences were made up on the basis of combinations of 3 infrequent glosses that were offered together on a slide to the signers.

Signs were selected that on November 17 occurred between 5 and 9 times in the corpus. A list of combinations of three glosses was created so that in the end each gloss would pass the threshold of 10 occurrences (including the instances that already occur in the corpus).

The set of stimuli is listed in the Appendix. Two native NGT signers created sentences or combinations of two sentences on the basis of the three words on the cue slide.

List of files

| Signer 1 (MvZ) | Stimuli | Duration |
|--------------------------|-------------------------|-----------------|
| SignSpeak_D1.8_MvZ_1.mov | Set 1-1 | 0:08:14 |
| SignSpeak_D1.8_MvZ_2.mov | Set 1-2 | 0:07:38 |
| SignSpeak_D1.8_MvZ_3.mov | Set 2-1 | 0:14:03 |
| SignSpeak_D1.8_MvZ_4.mov | Set 2-2 | 0:10:01 |
| SignSpeak_D1.8_MvZ_5.mov | Set 3-1 | 0:18:29 |
| SignSpeak_D1.8_MvZ_6.mov | Sets 1-3, 2-3, 3-2 | 0:34:07 |
| Signer 2 (YN) | | |
| SignSpeak_D1.8_YN_1.mov | Sets 1-1, 1-4, 2-4 | 0:18:01 |
| SignSpeak_D1.8_YN_2.mov | Sets 1-2, 2-1, 2-1, 3-1 | 0:25:26 |

Recording conditions

- Same Sony HDV camera as used for the Corpus NGT recordings.
- Same background as for the Corpus NGT recordings.
- Camera oriented straight at the signer (versus the slightly diagonal orientation in the Corpus NGT)

Tasks

- Create a sentence or max. a sequence of two sentences that minimally contains the three signs represented by glosses on the cue slide.

Appendix: Elicitation materials for the new sentences in D1.8

Each row presents the three glosses that were offered in combination as the assignment to the signers.

Set 1: Current frequency of occurrence of 5 glosses in the corpus

Set 2: Current frequency of occurrence of 6 or 7 glosses in the corpus

Set 3: Current frequency of occurrence of 8 or 9 glosses in the corpus

Set 1 and 2 contained 4 subsets with identical glosses, however, in different combinations of 3 glosses in each of the four lists in order to create multiple different sentences containing the same unique low frequency glosses.

Set 3 contained 2 subsets with identical glosses. We only used two lists here, given the already nearly sufficient rate of the glosses in the corpus NGT. The aim was to boost each of the low frequency glosses (occurring 5 to 9 times in the corpus) to a frequency of more than 10.

Set 1, subset 1

| | | |
|----------------------|---------------|------------|
| 1-UUR | EERLIJK | OMCIRKELEN |
| #T | REACTIE | PAKKEN-A |
| GOEDKOOP | ZIEK | STEL |
| BOUWMARKT=BOUW^MARKT | WAARDE | SPIJT |
| B | VERDER | NATUURLIJK |
| TWIJFEL | SELECTEREN | NIET-MOGEN |
| OPKOMST | PERSOON.PL | NIET-B |
| #DIALECT | FRIESLAND | BEGELEIDEN |
| MAG-NIET | VISICOM | KANT |
| NAAM | HOPEN | GEWELDIG |
| JAREN | VAST | JAAR-A |
| PLANT | NURNBERG | PROTEST |
| ETEN | #NEE | VROUW |
| OPVOEDEN | OMGEVING | INSTROMEN |
| DONDERDAG-A | KLEINE-DINGEN | INFORMEREN |
| MAANDAG-A | STERKTE | RICHTING |
| WOORD.PL | LIJKEN | IEMAND |
| BOE | OPLOSSING | BEZOEKEN |
| VISUEEL | KOPIEREN | TOE-MAAR |
| TOEVALLIG | DAT | WAARVOOR |
| BUITENLAND | ZIEN | LACHEN |

Set 1. Subset 2

| | | |
|----------------------|------------|------------|
| SELECTEREN | NIET-MOGEN | MAG-NIET |
| BEZOEKEN | JAAR-A | IEMAND |
| B | GOEDKOOP | STEL |
| BUITENLAND | RICHTING | JAREN |
| KANT | SPIJT | LACHEN |
| #NEE | TOEVALLIG | DAT |
| REACTIE | TWIJFEL | WAARVOOR |
| HOPEN | ETEN | PROTEST |
| BEGELEIDEN | OPVOEDEN | BOE |
| BOUWMARKT=BOUW^MARKT | OMGEVING | PERSOON.PL |
| VAST | PLANT | LIJKEN |

ZIEK
INSTROMEN
WAARDE
TOE-MAAR
VISUEEL
INFORMEREN
OPKOMST
OMCIRKELEN
VISICOM
NAAM

PAKKEN-A
#DIALECT
1-UUR
NURNBERG
WOORD.PL
KLEINE-DINGEN
EERLIJK
DONDERDAG-A
VROUW
KOPIEREN

NIET-B
STERKTE
OPLOSSING
GEWELDIG
FRIESLAND
ZIEN
VERDER
NATUURLIJK
MAANDAG-A
#T

Set 1, subset 3

#DIALECT
IEMAND
OPKOMST
NATUURLIJK
VROUW
DAT
SELECTEREN
REACTIE
NURNBERG
ZIEN
KOPIEREN
ETEN
VISICOM
BUITENLAND
VISUEEL
OMGEVING
OMCIRKELEN
NIET-B
MAG-NIET
HOPEN
BEZOEKEN

OPLOSSING
PLANT
FRIESLAND
MAANDAG-A
EERLIJK
WAARVOOR
DONDERDAG-A
NAAM
GOEDKOOP
BOUWMARKT=BOUW^MARKT RICTHING
#T
B
1-UUR
INSTROMEN
BEGELEIDEN
INFORMEREN
JAAR-A
NIET-MOGEN
JAREN
PAKKEN-A
KLEINE-DINGEN

WOORD.PL
LIJKEN
TOE-MAAR
ZIEK
STEL
GEWELDIG
OPVOEDEN
STERKTE
TOEVALLIG
LACHEN
#NEE
KANT
PROTEST
TWIJFEL
BOE
WAARDE
VERDER
PERSOON.PL
VAST
SPIJT

Set 1, subset 4:

MAG-NIET
PERSOON.PL
PLANT
KLEINE-DINGEN
NURNBERG
OPKOMST
IEMAND
1-UUR
TOEVALLIG
BEGELEIDEN
ETEN
RICHTING
WOORD.PL
LIJKEN
TWIJFEL
SELECTEREN
VISICOM
INSTROMEN
EERLIJK
OMGEVING

KOPIEREN
NIET-B
ZIEN
NAAM
OPVOEDEN
JAREN
STERKTE
BEZOEKEN
BUITENLAND
NATUURLIJK
#NEE
INFORMEREN
ZIEK
TOE-MAAR
BOUWMARKT=BOUW^MARKT GOEDKOOP
WAARVOOR
OPLOSSING
MAANDAG-A
VISUEEL
LACHEN

SPIJT
#DIALECT
JAAR-A
VROUW
DAT
B
DONDERDAG-A
OMCIRKELEN
GEWELDIG
BOE
PROTEST
VERDER
STEL
WAARDE
REACTIE
#T
NIET-MOGEN
VAST
KANT

FRIESLAND

PAKKEN-A

HOPEN

Set 2, subset 1:

BLAUW
BESTAAN
MET
JONGEREN
3
GELUID
AFLEIDEN
WONEN-A
#M
BESTEMPELEN-ALS
CI-IMPLANTEREN
VOOR-A
#L
KIND
PAD
ZIJN
LIMBURG
TWIJFELEN
BRABANT
#IVF
SUBTIEL
SORRY
DRINKEN
ZOMAAR
KOMMA
ONDER
GEBARENTOLK=GEBAREN-A^TOLK
ENZOVOORT
VANZELF
WERVEN
IDEE
#AAN
LANGZAAM
WEET
HOMO
NATUUR
WETEN-1
HOEVEN
ZINNEN
MACHT
ZUS-E
STAAN
PAPA
BREED
EENZAAM

VIERKANT-OOR
RECHT
DVD
WEST
AANPASSING
ONDERZOEK
KEUZE
BEWIJS
DUITSLAND
TUSSEN-HAAKJES
ELKAAR
DING.PL
PAS-B
VERWIJDEREN
SEKS
LUKKEN
EFFATHA
14
STEM-B
GROEIEN
TREIN
SPANJE
WANDELEN
ANDERE-KANT
WELKE
TELLEN
JAPAN
#N
AMSTEL
BEKIJKEN
#B
#S
BEPERKEN
MAKEN-B
SIMULTAAN
VREEMD-B
KORT
ORGANISEREN
PATIENT
ZAND
SLECHT
VRIEND-B
ZONDER
RUSTIG

VERGETEN
SCHOOL
JAAR-C
POSITIEF
DOKTER
HARD
BEVESTIGEN
#TE
ANDERSOM
VASTSTELLEN
EISEN
SAAI
ONTDEKKEN
AFSPREKEN
13
VOOR-C
BIJEENKOMEN
VADER-B
RAAR
NA
VERANTWOORDELIJK
HANDSUP
IDENTITEIT
ROTTERDAM
ALLES
VERHUIZEN
KAARTEN #R
VEEL
SMS
BEGINNEN
STEEN
STEM-A
#H
ZELFDE-C
ZATERDAG-A
STUDIE
SPELEN
COMBINATIE
SYNONIEM
VERSCHILLEND
ERGENS
BETEKENIS
EMPOWERMENT
50

Set 2, subset 2

VERSCHILLEND
DVD
EENZAAM
SORRY
#R

WERVEN
BEKIJKEN
#L
WETEN-1
JAAR-C

PAPA
VERGETEN
ELKAAR
MAKEN-B
KEUZE

| | | |
|-----------------|----------------------------|------------------|
| ZIJN | LUKKEN | #B |
| RAAR | COMBINATIE | VANZELF |
| KIND | ONTDEKKEN | VERANTWOORDELIJK |
| BESTAAN | DING.PL | WEET |
| CI-IMPLANTEREN | MACHT | ORGANISEREN |
| SCHOOL | SAAI | IDEE |
| HOMO | ANDERSOM | SEKS |
| LANGZAAM | STEM-A | BEGINNEN |
| VOOR-A | #N | NA |
| #AAN | ANDERE-KANT | JONGEREN |
| HOEVEN | EMPOWERMENT | ONDER |
| SIMULTAAN | BIJEENKOMEN | ZAND |
| TREIN | BEVESTIGEN | WELKE |
| ALLES | GROEIEN | ZONDER |
| GELUID | ONDERZOEK | TUSSEN-HAAKJES |
| #S | HARD | AANPASSING |
| STEEN | PATIENT | 50 |
| STUDIE | BEPERKEN | LIMBURG |
| VOOR-C | TWIJFELEN | DUITSLAND |
| VREEMD-B | VERWIJDEREN | WANDELEN |
| SMS | #M | STAAN |
| VIERKANT-OOR | NATUUR | KORT |
| SYNONIEM | PAS-B | POSITIEF |
| BRABANT | RUSTIG | HANDSUP |
| ERGENS | RECHT | BREED |
| VRIEND-B | GEBARENTOLK=GEBAREN-A^TOLK | #TE |
| BESTEMPELEN-ALS | SPELEN | BLAUW |
| JAPAN | ZELFDE-C | AMSTEL |
| ZOMAAR | WONEN-A | SPANJE |
| EFFATHA | ZATERDAG-A | AFSPREKEN |
| BETEKENIS | MET | 3 |
| 13 | ZUS-E | VERHUIZEN |
| 14 | KAARTEN | DOKTER |
| #H | VASTSTELLEN | TELLEN |
| BEWIJS | #IVF | DRINKEN |
| IDENTITEIT | KOMMA | VADER-B |
| EISEN | STEM-B | ENZOVOORT |
| VEEL | SUBTIEL | PAD |
| SLECHT | WEST | ZINNEN |
| AFLEIDEN | ROTTERDAM | |

Set 2, subset 3

| | | |
|----------------------------|------------------|------------|
| RECHT | 14 | SMS |
| PAS-B | SEKS | BEVESTIGEN |
| RUSTIG | VANZELF | ALLES |
| ANDERE-KANT | BEKIJKEN | WERVEN |
| ZOMAAR | DUITSLAND | IDENTITEIT |
| ZONDER | MET | ZAND |
| TWIJFELEN | BESTAAN | SPANJE |
| BLAUW | STUDIE | STAAN |
| ELKAAR | DVD | WANDELEN |
| GEBARENTOLK=GEBAREN-A^TOLK | | BREED SAAI |
| ANDERSOM | VERANTWOORDELIJK | JAAR-C |
| VOOR-C | HANDSUP | WELKE |
| VADER-B | #IVF | NA |
| ONTDEKKEN | WETEN-1 | SIMULTAAN |

JONGEREN
LANGZAAM
SYNONIEM
BRABANT
ZUS-E
MACHT
DING.PL
BEGINNEN
ORGANISEREN
KORT
VERHUIZEN
#L
BESTEMPELEN-ALS
VASTSTELLEN
#H
ZIJN
IDEE
BETEKENIS
#B
13
ENZOVOORT
COMBINATIE
KEUZE
SLECHT
#AAN
HARD
VERGETEN
SUBTIEL
HOEVEN
50
VERWIJDEREN

VEEL
AMSTEL
#S
#TE
KAARTEN
LUKKEN
PATIENT
STEM-A
ONDER
ZINNEN
AANPASSING
ZELFDE-C
POSITIEF
RAAR
TREIN
GELUID
#N
VOOR-A
TELLEN
STEM-B
EMPOWERMENT
DOKTER
ROTTERDAM
AFSPREKEN
PAPA
BIJEENKOMEN
BEWIJS
AFLEIDEN
KIND
GROEIEN
DRINKEN

SPELEN
EISEN
VREEMD-B
WEST
WONEN-A
NATUUR
STEEN
VRIEND-B
JAPAN
EFFATHA
#R
PAD
ONDERZOEK
ZATERDAG-A
SORRY
ERGENS
EENZAAM
TUSSEN-HAAKJES
3
LIMBURG
SCHOOL
KOMMA
HOMO
#M
CI-IMPLANTEREN
VIERKANT-OOR
WEET
VERSCHILLEND
BEPERKEN
MAKEN-B

Set 2, subset 4

COMBINATIE
BLAUW
SYNONIEM
AMSTEL
SIMULTAAN
DOKTER
KIND
VERANTWOORDELIJK
BEKIJKEN
ZELFDE-C
WANDELEN
SEKS
SAAI
VOOR-A
ZAND
LUKKEN
AFLEIDEN
STEM-A
SCHOOL
LIMBURG
HOMO
POSITIEF
SUBTIEL

VANZELF
ZIJN
JAPAN
PAD
KAARTEN
STUDIE
MACHT
KORT
VADER-B
#L
BIJEENKOMEN
ONDERZOEK
AFSPREKEN
ROTTERDAM
JAAR-C
SPANJE
13
#S
DUITSLAND
VERGETEN
#B
PATIENT
SLECHT

ZUS-E
VEEL
STEEN
NATUUR
VREEMD-B
ERGENS
WETEN-1
KEUZE
ONTDEKKEN
STEM-B
PAPA
SPELEN
RECHT
SORRY
BETEKENIS
BEPERKEN
ZINNEN
ZONDER
BRABANT
LANGZAAM
BEGINNEN
RUSTIG
EISEN

#IVF
JONGEREN
VOOR-C
DRINKEN
WELKE
EFFATHA
14
ANDERSOM
SMS
#H
MET
BEWIJS
#TE
VIERKANT-OOR
DING.PL
ONDER
GROEIEN
ALLES
A^TOLK
STAAN
WERVEN
VERSCHILLEND
BREED

#AAN
GELUID
PAS-B
TREIN
BESTEMPELEN-ALS
MAKEN-B
3
RAAR
DVD
IDENTITEIT
WEST
ENZOVOORT
VERWIJDEREN
50
WONEN-A
AANPASSING
HANDSUP
HOEVEN

KOMMA
VERHUIZEN
BEVESTIGEN
EENZAAM

TWIJFELEN
ORGANISEREN
ZOMAAR
ANDERE-KANT
BESTAAN
VASTSTELLEN
#R
EMPOWERMENT
WEET
ELKAAR
#N
CI-IMPLANTEREN
#M
IDEE
NA
HARD
VRIEND-B
GEBARENTOLK=GEBAREN-

TELLEN
ZATERDAG-A
TUSSEN-HAAKJES

Set 3, subset 1

VOORUIT
ZIEKENHUIS=ZIEK^HUIS-A
DISCUSSIEREN
60
JAMMER
#D
VERSPREID
WEEK
3-B
WAAR-A
DENKEN-1
#MSN
HOUDEN
STIMULEREN
OMA
VISIE
LIJST
CENTRUM
MAMA
ERVAREN
GELOVEN
TESTEN
IX:P
ZELFDE
OPVALLEN
NEE-B
HOND-A

LOSLATEN
BETALEN
LOPEN
BLIJKEN
#ASL
AANBOD
OPNIEUW
BESTEMPELEN
ZIN
TOLKEN
OPA
ONDERWERP
COMPUTER
BIJ
SEXUOLOOG
MIX
ENGELS
BEELD
ONTSPANNEN
LIEVER
MOGELIJK
SCHRIKKEN
ONDERDANIG
VOETBALLEN
ZITTEN
GELUKKIG
KANS

POLITIEK
OFFICIEEL
SOCIAAL
12
TYPEN
ZOIETS
OVER-C
VERWACHTEN
HULPMIDDEL=HELPEN-A^HELFT
COLLEGA
LEKKER
EHM
WINKEL
#G
OVEREENKOMST
ORGANISATIE
PRAKTIJK
MIS
MATIG
GEBOREN-WORDEN
NIEUW
OVER-B
MENS.PL
VERSTAND
TWEEN:ZIJ
20
GAUW

Set 3, subset 2

SOCIAAL

3-B

OVEREENKOMST

| | | |
|---------------------------------|-------------|------------|
| SCHRIKKEN | OMA | MIX |
| DENKEN-1 | ZITTEN | OFFICIEEL |
| GEBOREN-WORDEN | MOGELIJK | BIJ |
| HULPMIDDEL=HELPEN-A^HELFTOVER-C | GELOVEN | HOND-A |
| MATIG | WAAR-A | OPVALLEN |
| ZIEKENHUIS=ZIEK^HUIS-A | BESTEMPELEN | ONDERWERP |
| BETALEN | KANS | MIS |
| LOSLATEN | MAMA | POLITIEK |
| STIMULEREN | COMPUTER | 20 |
| DISCUSSIEREN | VERSTAND | GELUKKIG |
| #ASL | OPA | ONDERDANIG |
| OVER-B | SEXUOLOOG | LIJST |
| ORGANISATIE | LOPEN | 60 |
| #G | BLIJKEN | VISIE |
| HOUDEN | VOORUIT | EHM |
| BEELD | TYPEN | GAUW |
| NEE-B | ZIN | MENS.PL |
| ZELFDE | PRAKTIJK | ERVAREN |
| WEEK | TWEEEN:ZIJ | LIEVER |
| AANBOD | VERSPREID | WINKEL |
| VERWACHTEN | ZOIETS | ONTSPANNEN |
| TOLKEN | COLLEGA | NIEUW |
| 12 | TESTEN | #MSN |
| #D | LEKKER | VOETBALLEN |
| ENGELS | JAMMER | IX:P |
| CENTRUM | | OPNIEUW |