

Vi har undersökt hur variation i prosodiska särdrag på syntetiserade ettordsyttranden uppfattas och påverkar lyssnarens respons. I en första studie bedömde försökspersoner vilken av tre givna tolkningar som passade bäst på syntetiserade renderingar av enstaviga färgord. Det visade sig att prosodi kan signalera verifiering på flera nivåer.

Vi nu gjort en ny serie försök där avsikten är att utröna om våra prosodiska mönster fungerar som väntat i en dialogsituation mellan människa och dator. Resultaten visas här i korthet.

Exempelmiljö: Higgins navigering för fotgängare

S: beskriv vad du ser  
U: jag ser ett brunt hus till vänster  
S: röd?  
U: nej... brunt

## Fel och verifiering i dialog

- Dialog är sällan "felfri"
- fel-detektering involverar ofta implicit (ovan och nedan) eller explicit (nedan) verifiering.

U [...] till höger ser jag ett rött hus.  
S (låg konfidens) Sa du 'ett rött hus'?  
S (hög konfidens) Ett rött hus... ok, ta till vänster [...?]

- Formulerade som fullständiga fraser
- Uppfattas ofta som tradiga
- Verifierar hela användaryttranden

## Verifiering med fragment

U [...] till höger ser jag en röd byggnad.  
S röd(?)

- Snabbt
- Kokuserar på problematiska fragment
- Används ofta i människa-människa-dialog

## Problem

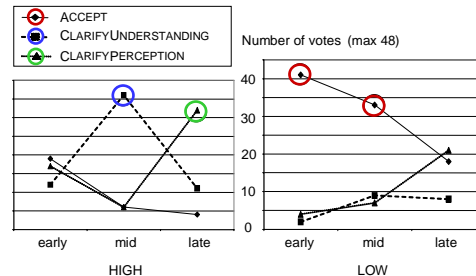
- verifiering med ettordsuttryck är potentiell tvetydigt
- Lite stöd av syntax och struktur
- Prosodi mer kritiskt
- Hur påverkar prosodin tolkningen av dessa yttranden?**

## Förståelse och verifiering på olika nivåer

- Allwood et al. (1992), Clark (1996)

Nivå		Verifiering
Acceptans	A accepterar vad B säger	OK, röd.
Förståelse	A förstår vad B säger	Menar du verkligen röd?
Perception	A hör vad B säger	Sa du röd?
Kontakt	A hör att B talar	

## Experiment I - resultatsammanfattning



- Tydlig skillnad mellan verifiering av perception och förståelse
- Inga effekter för person, färgord, eller duration (!)
- Tre tydliga prototyper
  - Acceptans – tidig låg topp
  - Förståelse – hög topp i mitten
  - Perception – sen hög topp

## Experiment II - design

- 8 försökspersoner fick till uppgift att tillsammans med en talande dator komma överrens om en färgmodell genom att berätta vilken av färgerna röd, gul, grön och blå som låg närmast en serie mellanliggande färger (orange, gulgrön, turkos respektive lila). De två alternativen visades till höger och vänster om den problematiska färgen i en slags triptyk, som den i bakgrunden här.
- Ganska svår uppgift där systemet kan ha anledning att bli konfunderat eller vara oense med försökspersonen.

S: Och den här?  
U: Blå.  
S: Blå?  
U: ...nej, grön!  
S: Grön.  
U: Ja.

## Fyra versioner av respons på färgangivelser

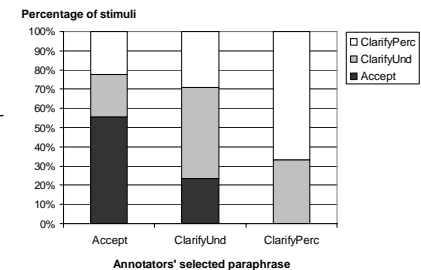
- tre repetitioner av färgordet enligt experimient I
  - accepterande
  - klarifierande av förståelse
  - klarifierande av perception
- ett neutralt accepterande i form av /m/ eller /a/

En Wizard kände igen färg-orden



## Experiment II - resultat 1

- 2 bedömare lyssnade till alla användarresponsen som följde direkt på en ettordsverifiering och försökte avgöra vilken av parafaserna som var mest trolig som föregående yttrande
- Bedömarna överens i 50% av fallen
- Bedömningarna överensstämmer med de hypotetiserade tolkningarna av de systemyttranden som faktiskt föregick användarresponserna



## Experiment II - resultat 2

Stimuli	Responstid
Acceptans	591 ms
Förståelse	976 ms
Perception	634 ms

- Tiden mellan ettordsverifiering och användarrespons mättes för alla instanser med hjälp av /naillon/ - programvara för talanalys
- Hypotestestning: acceptans är mindre belastande än osäkerhet kring vad man hört, som är mindre belastande än oförståelse
  - Acceptans: kort responstid
  - Perceptionsverifiering: längre responstid
  - Verifiering av förståelse: längst responstid
- Resultatet stöder hypotesen ('one-way repeated measures ANNOVA')