

Frank Kügler

# Multimodal marking of prominence in communication

73. StuTS 27.05.2023 – Goethe University Frankfurt



# What is prominence in communication?

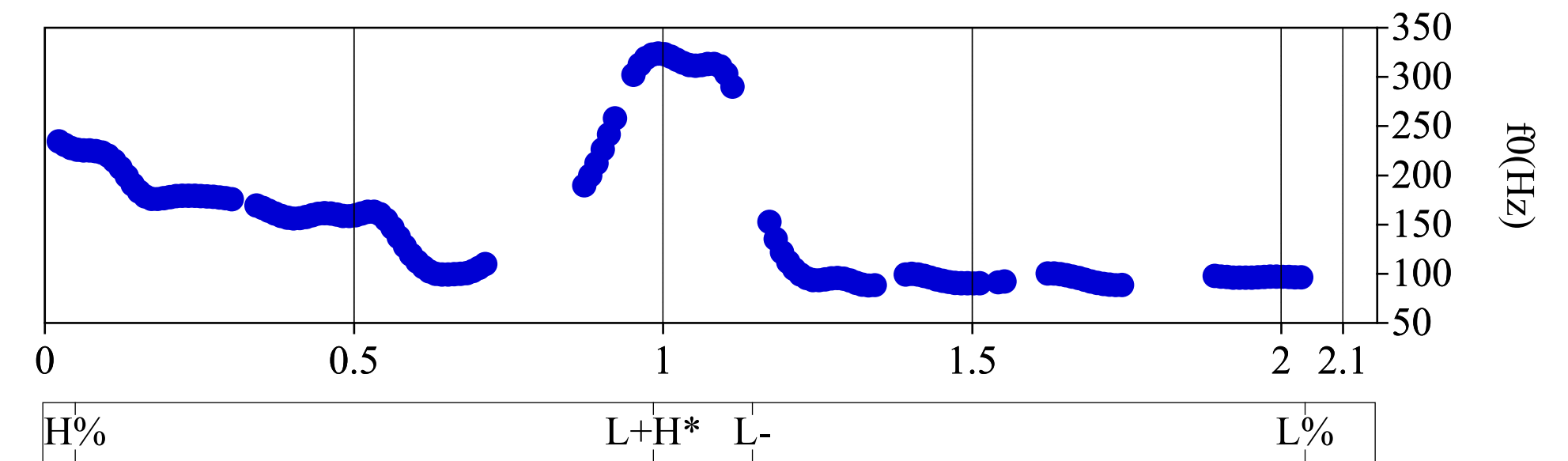
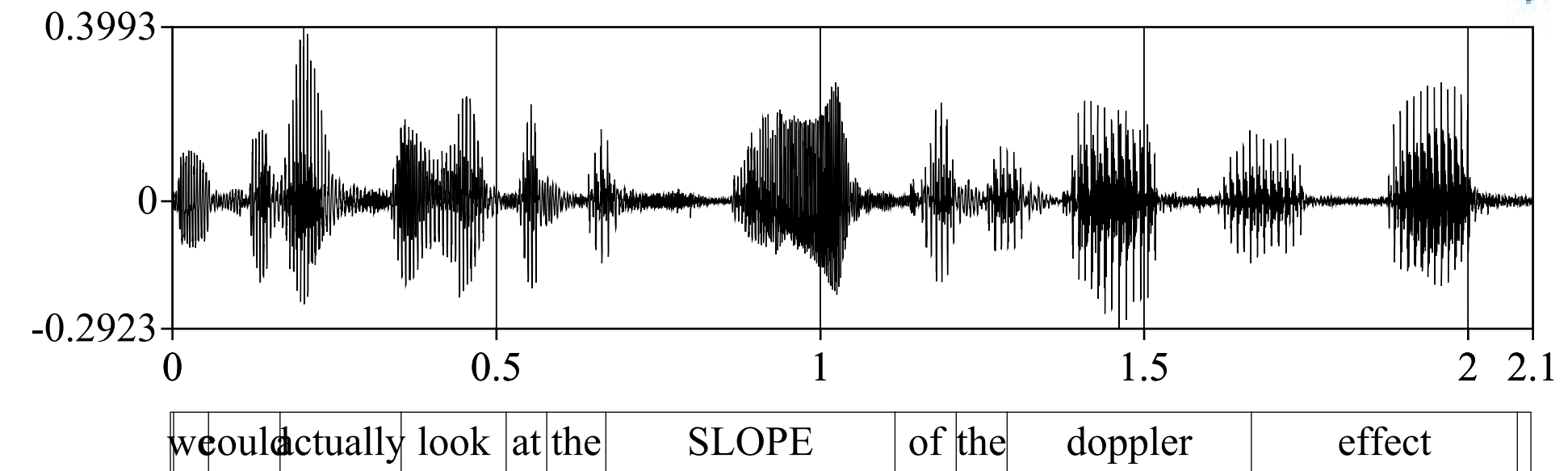


(Data from Rohrer et al. 2020)

# Multimodal speech

The signal contains different levels.

- Sequence of words as sound
  
  - Speech melody – intonation
  
  - Prosody – a mode of language (parallel to words)
  
  - Visual channel, movement of the hand / thumb and pointing finger
  
  - The visual channel is another mode of language
- ‘gestures and speech are two modalities of the same framework’ (McNeill 1992:23)



# Roadmap of the talk

## Background

- Prominence in Linguistics
- Multimodality of Speech
- Prosody, Co-Speech Gestures, Prosody-Gesture-Link

## Corpus Study (SaGA Corpus)

## Experimental study on prominence marking (work in progress)

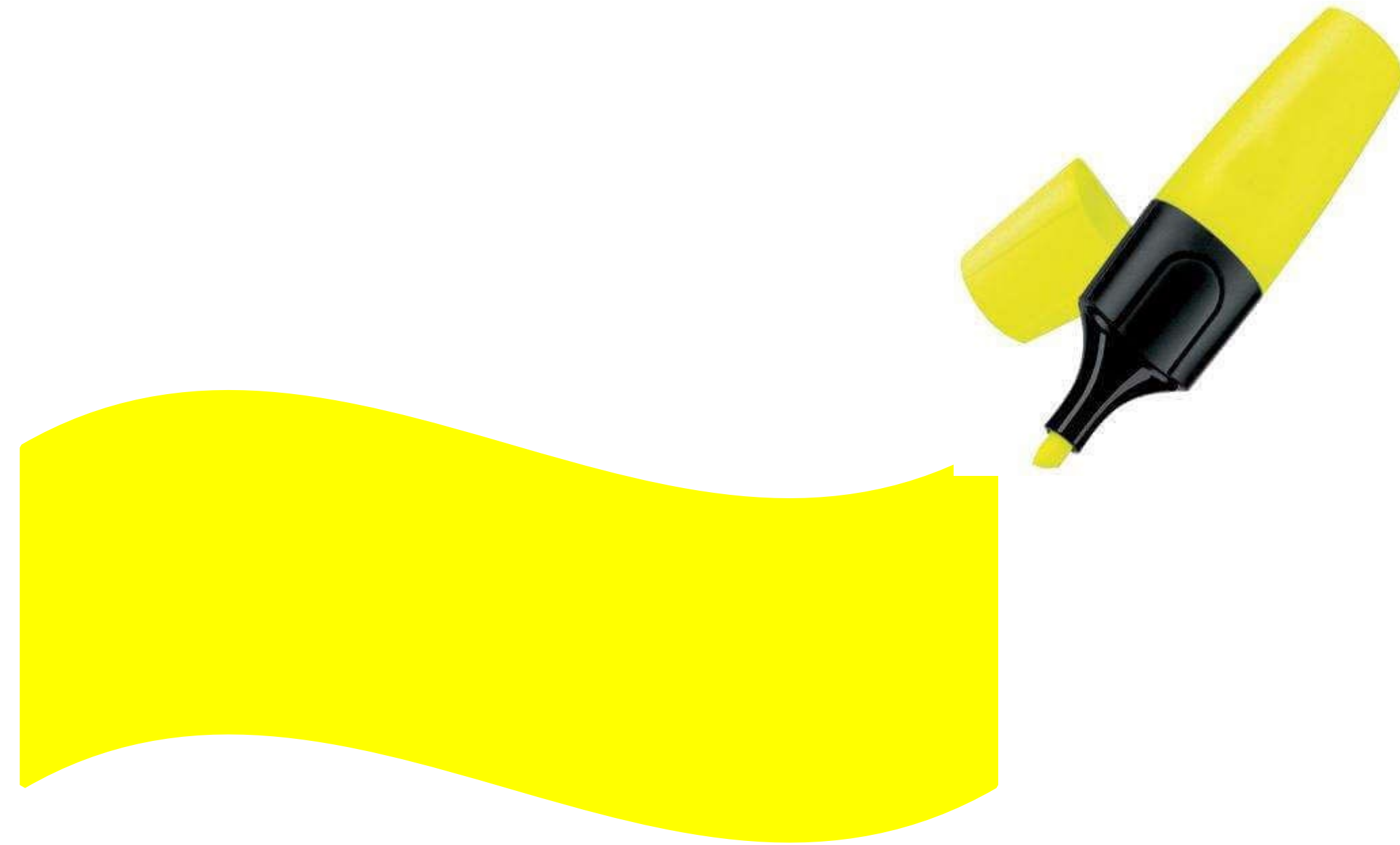
## Results

- Synchronisation of Gestures with Prosody
- Factor Prominence
- Overlapping Functions of Gestures

## Discussion

**Prominence attracts prosodic and visual marking –  
the more pragmatic prominence the more prominent are acoustic and gestural cues**

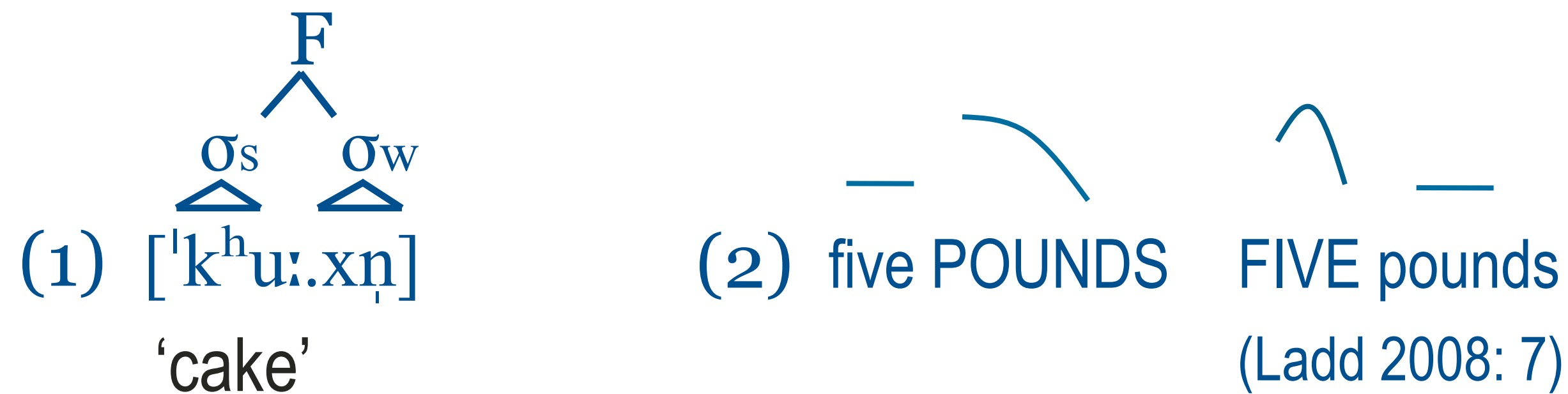
# Prominence in Linguistics



# Prominence in Linguistics

“Prominence as a **relational property**. [...] Prominence entails the property of being an **‘organizational principle’** for **linguistic structure** [...] or, in other words, that a **prominent element** organizes its environment, **providing a structure** for the context in which it appears [...].”

(Grice & Kügler 2021: 253)



**Relation** – more or less prominent elements  
**Elements** – linguistic constituents  
**Prosodic prominence** – prosodic structure and constituents

- Alternation of strong – weak elements  
 (stressed / unstressed syllables)  
 (accented / unaccented words)

...

# Prosody

What is prosody about? (Gussenhoven & Chen 2020)

- Form: phonetic cues as pitch, intensity, duration, spectral patterns
- Functions: How cues create communicative effects
- A more recent perspective is, how these cues signal prosodic structure.

Prosody:

“core prosodic elements [are] tone, stress, prosodic constituents, and intonation.”

(Gussenhoven & Chen 2020:4)

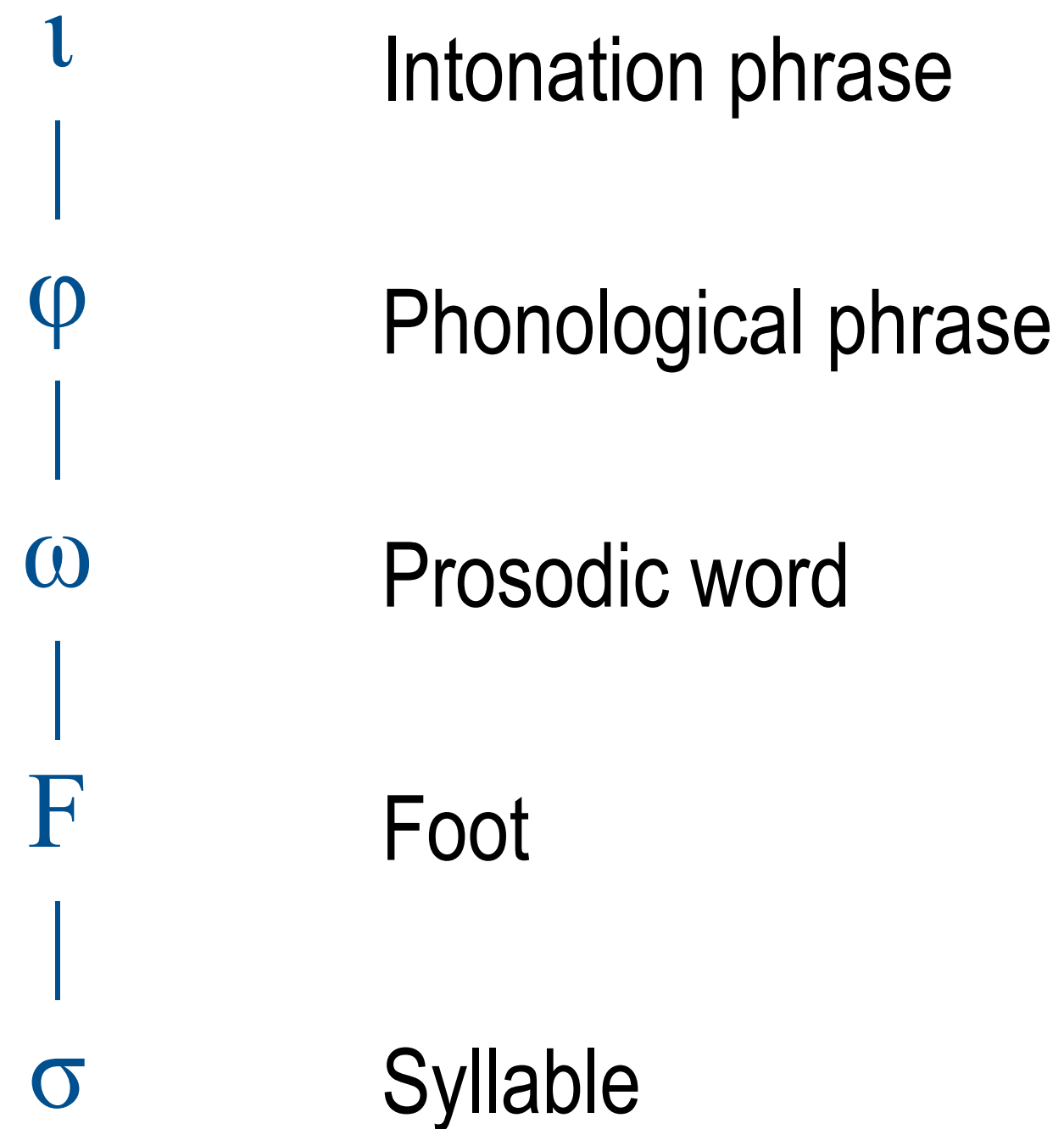
Intonation itself needs a definition:

“The use of *suprasegmental* phonetic features to convey ‘postlexical’ or *sentence-level* pragmatic meaning in a *linguistically structured way*” (Ladd 2008:4)

Distinction between word-level and phrase-level prosodic properties.

# Prosodic structure

## The prosodic hierarchy –



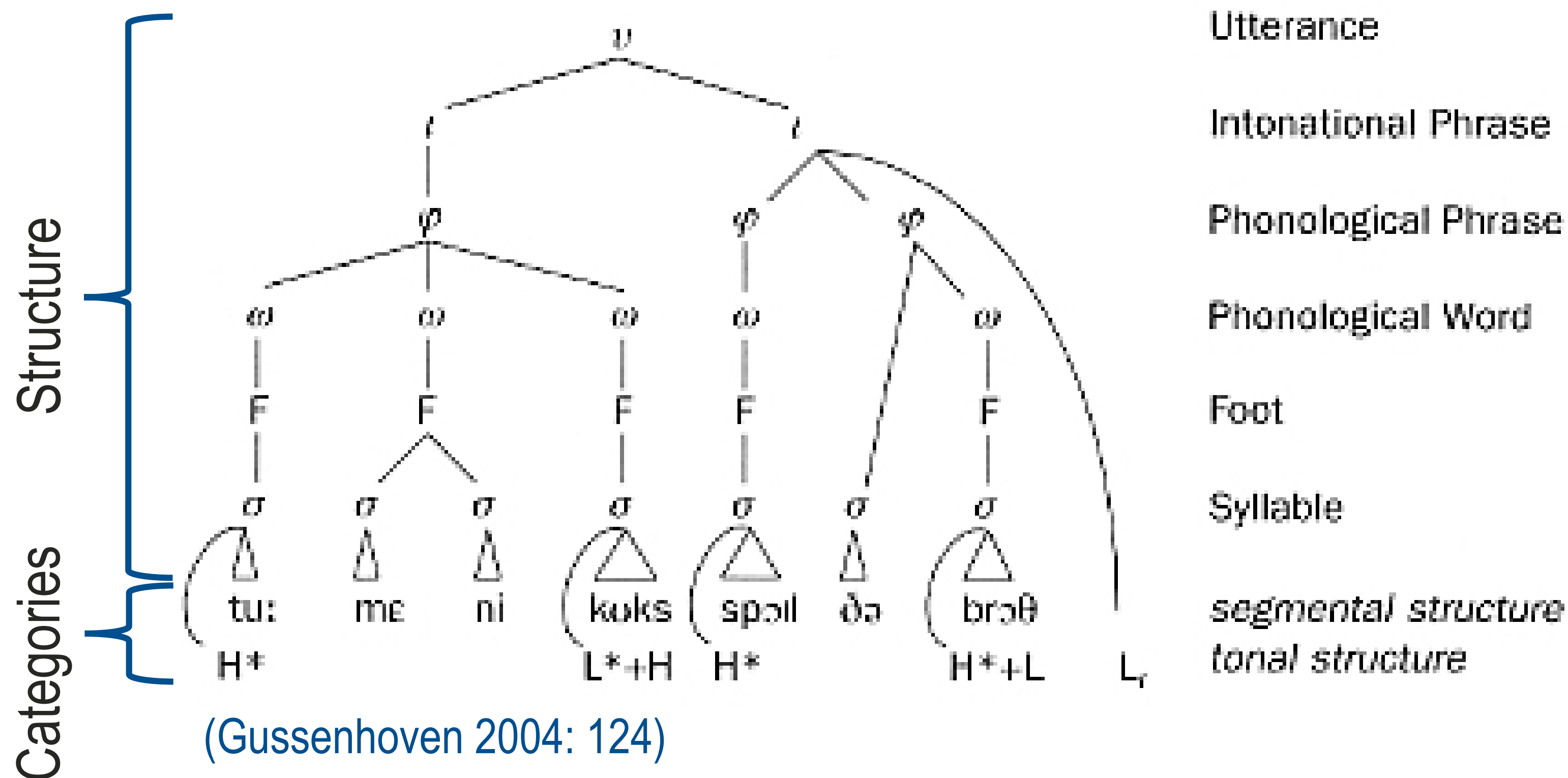
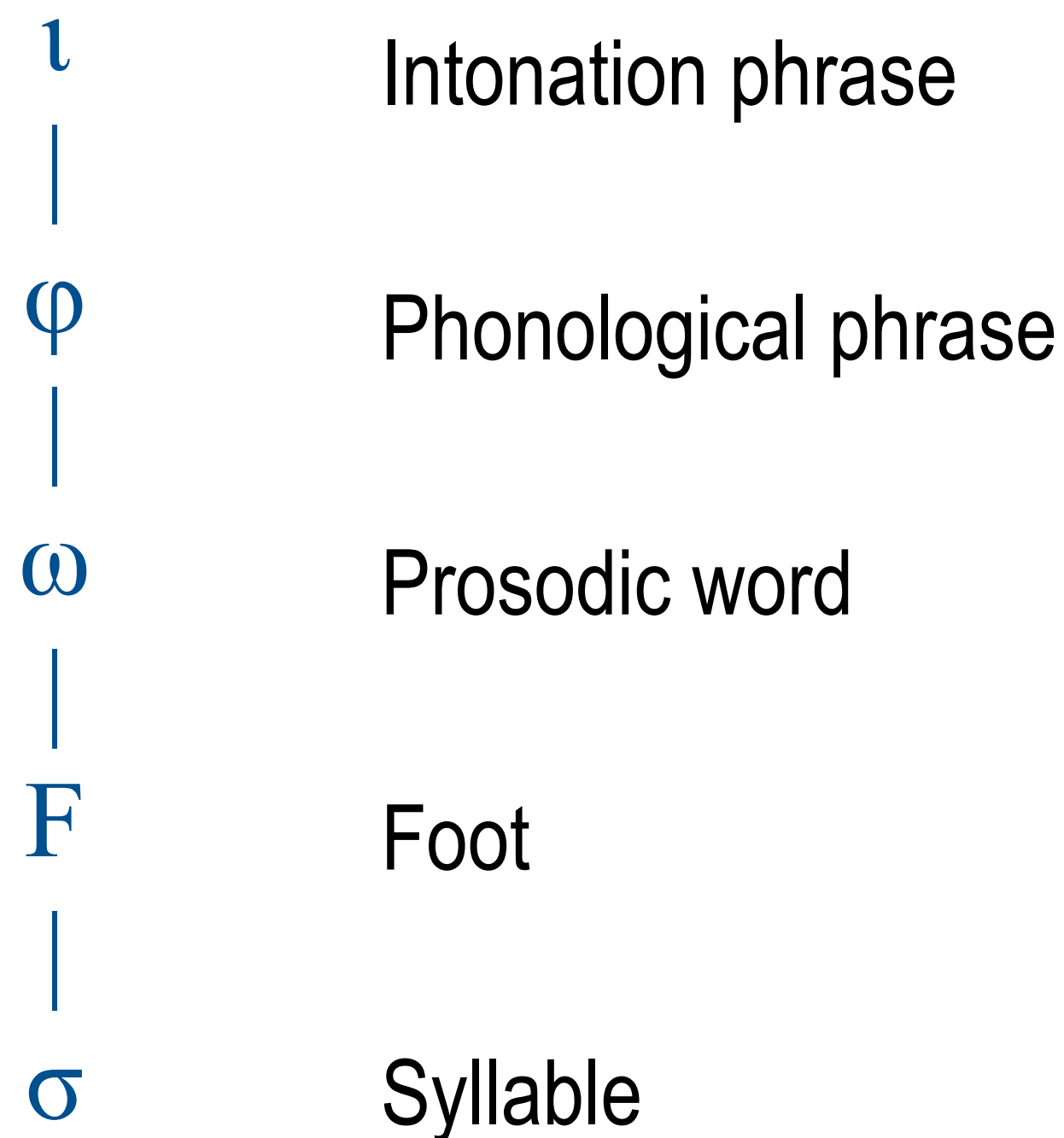
The most common  
prosodic constituents

(Nespor & Vogel 1986; Selkirk 1984;  
most recently, Féry 2017)



# Prosodic structure

## The prosodic hierarchy –

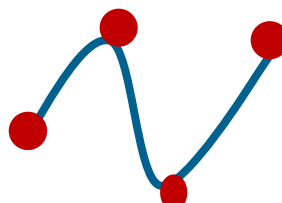



(Nespor & Vogel 1986; Selkirk 1984;  
most recently, Féry 2017)

# Linearity of tonal structure

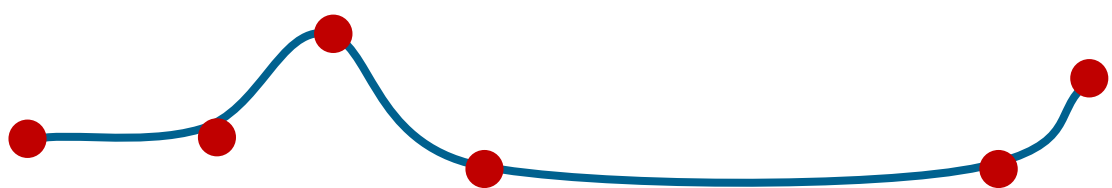

(3) A: I hear Sue's taking a course to become a driving instructor.

B: Sue!?

(4) A: I hear Sue's taking a course to become a driving instructor.

B: A driving instructor!?

(Ladd 1996: 44)

“AM theory represents a pitch contour phonologically as sequences of discrete intonational events.”

(Ladd 1996:43)

→ Two types of tonal events – pitch accents & boundary tones.

→ Internal structure of a pitch contour – events & transitions.

# Pitch accents

Definition of pitch accent (Ladd 1996:45f)

“a local feature of a pitch contour – usually but not invariably a *pitch change*, and often involving a local maximum or minimum – which signals that the syllable with which it is associated is *prominent* in the utterance.”

- Independence of Tones and Text
- Different types of Pitch Accents (H\*, L\*, H\*L, ...)
- Relative metrical prominence relations

(5) [Marianna]



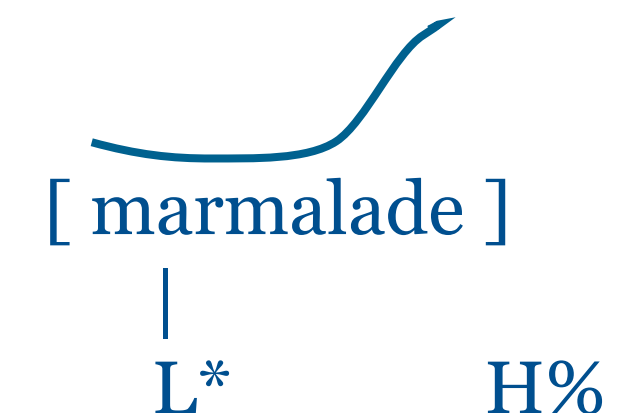
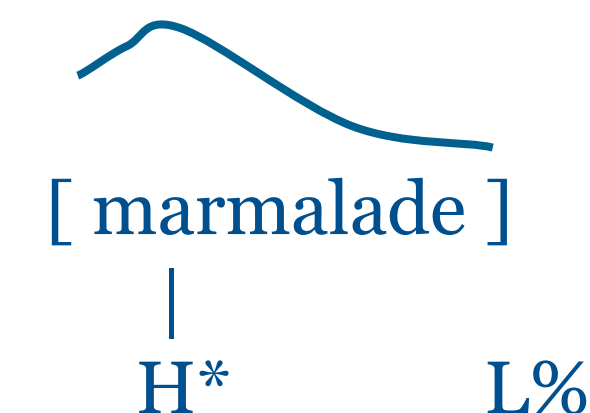
# Boundary tones – Where do they occur?

A boundary tone is associated with the very end of an intonation phrase.

- Recall: Pitch accents associate with metrically strong syllables (Pierrehumbert, 1980; Gussenhoven, 2004; Ladd, 2008)
- A boundary tone: F0 movement at an edge adjacent element
- Intervening unstressed syllables between a pitch accent and a phrase boundary show the independence of boundary tones from pitch accents.
- Metrically independent – realization independent of stress



(6) [Marianna made the marmalade?] (Pierrehumbert, 1980)



# Summary – Prosodic categories

## Prosodic categories (phrase level)

Pitch accents (head of a prosodic domain)

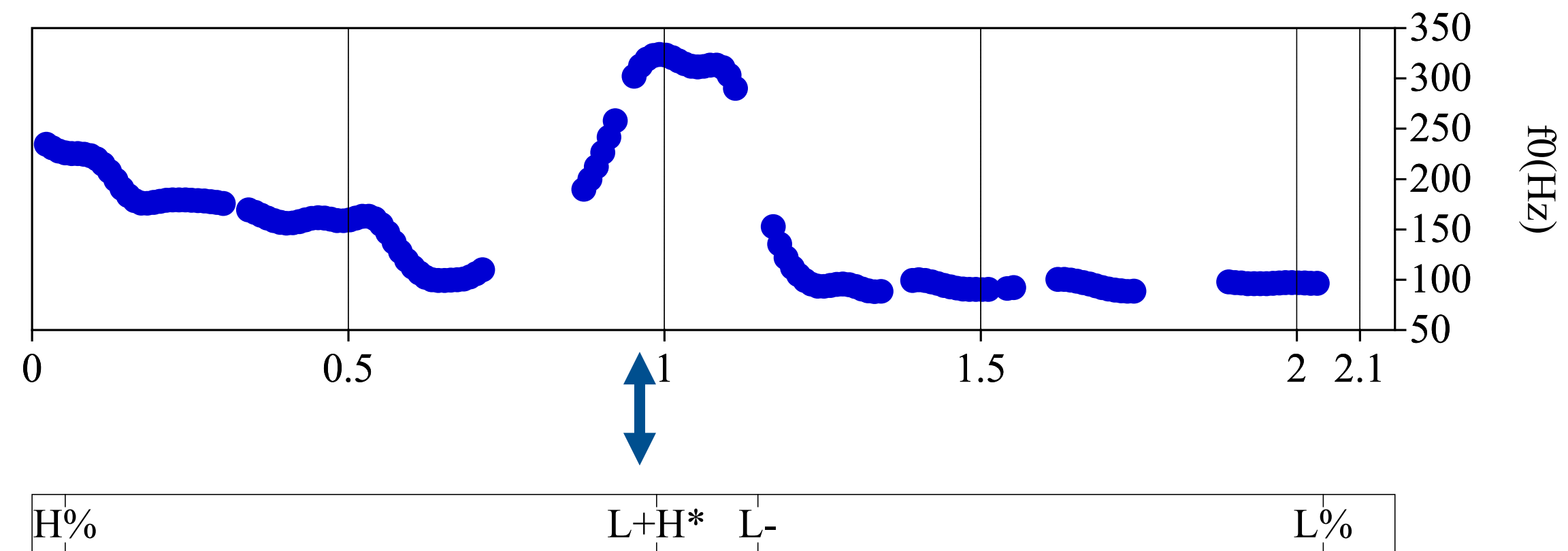
### highlighting function

➤ prominence

Boundary tones

### demarcation function

➤ structure











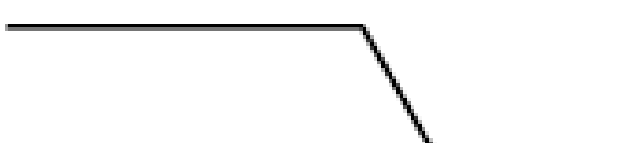

# Intonation – what do the categories achieve in perception?

## Linguistic function of prosody – Phrasing

How many people came for a visit?

Paul, wer war gestern bei dir zu Besuch?

a.      
 a. meine Schwester Anastasia und Angelique  
 $H^*LH\%$   $H^*LH\%$   $H^*\rightarrow !H^*LL\%$  

b.      
 b. meine Schwester Anastasia und Angelique  
 $H^* H\%$   $H^*LH\%$   $H^*\rightarrow !H^*LL\%$  

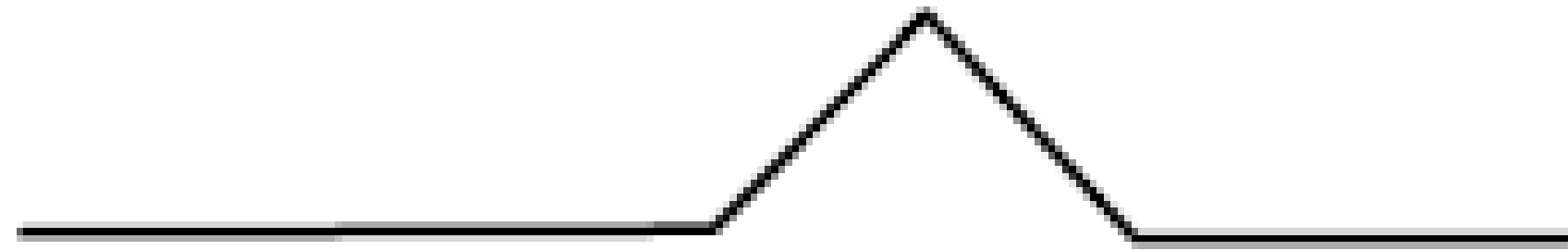
(Peters 2014)

# Intonation – what do the categories achieve in perception?

## Linguistic function of prosody – **Highlighting**



(1) a. sie lebt in Oldenburg



b. sie lebt in Osnabrück

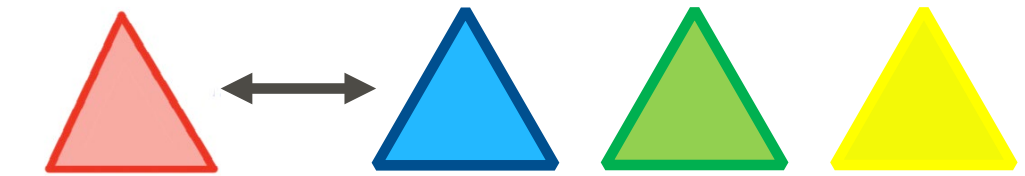
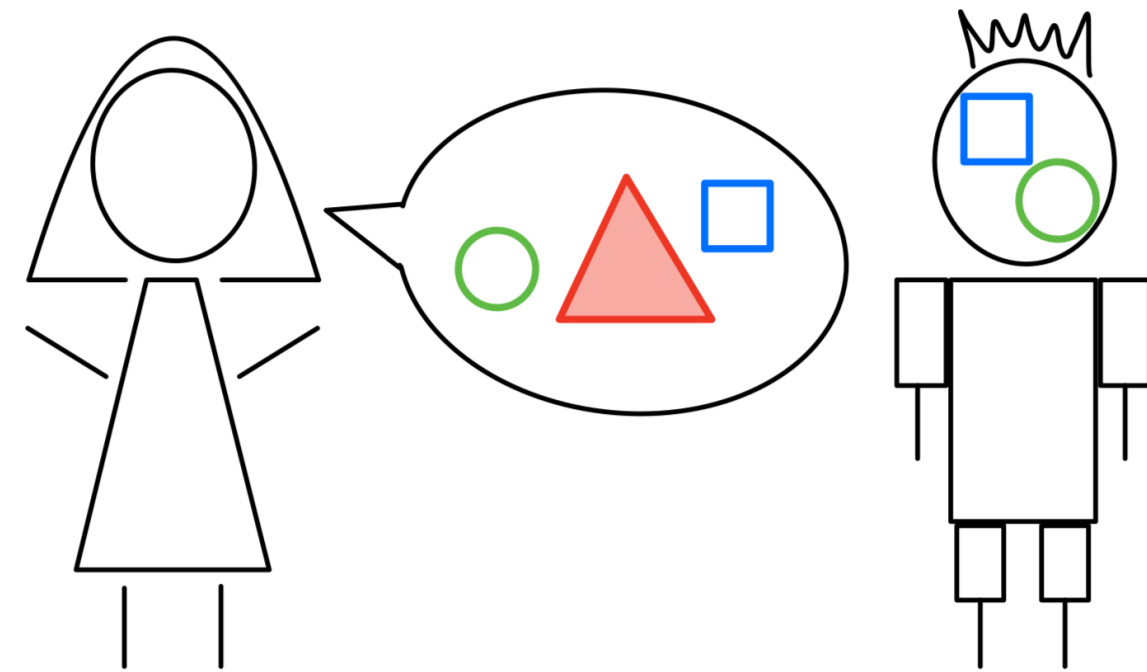


(Peters 2014)

# Prominence and Focus

## Highlighting = Focus (information structure)

“Focus indicates the presence of alternatives that are relevant for the interpretation of linguistic expressions.” (Krifka 2008: 247)



### Focus

Presence of alternatives in the discourse

“... we could actually look at the **SLOPE** of *the Dopplar effect*.”

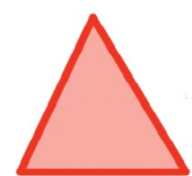


**Focus**



*background*

The focus is prominent. The background is less prominent.





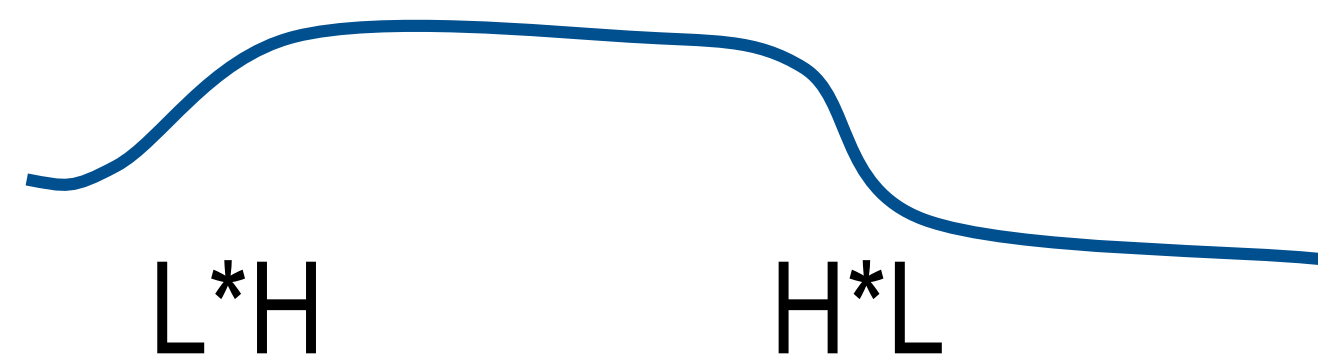
# Prosody, Prominence and Focus

Pitch accents can be used as indicators for prominence (Ladd 2008).

- Higher  $f_0$  → higher prominence (Baumann & Röhr 2015, Kügler & Calhoun 2020).
- Higher prosodic prominence indicates greater newness or informativeness

(7) (a)

Erzähl mir bitte, was passiert ist. 📢

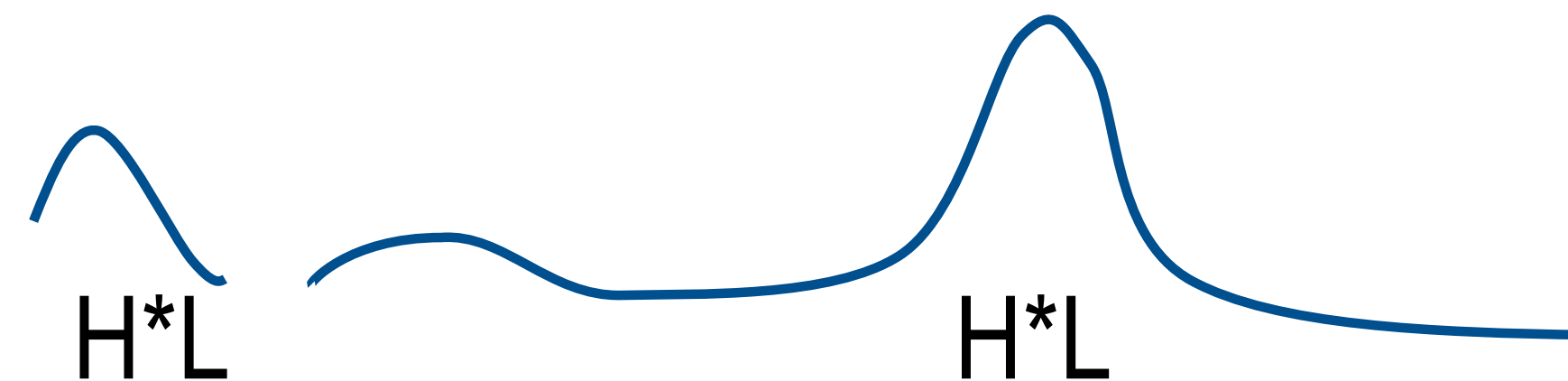


[Martin hat den Wal gesehen.]<sub>foc</sub>

‘Martin has seen the whale.’

(b)

Hat Martin den Frosch gesehen? 📢



Nein, Martin hat den **[Wal]**<sub>foc</sub> gesehen.

‘No, Martin has seen the whale.’

(Kügler & Gollrad 2015)

(8) lila – lila 📢

narrow – corrective focus → degrees of prominence (Data from MultiS)

## Co-speech gestures



# Co-speech gestures

Gestural research is a “new” discipline in linguistics  
(Gregori et al. to appear)

→ contributes to **(visual) communication**

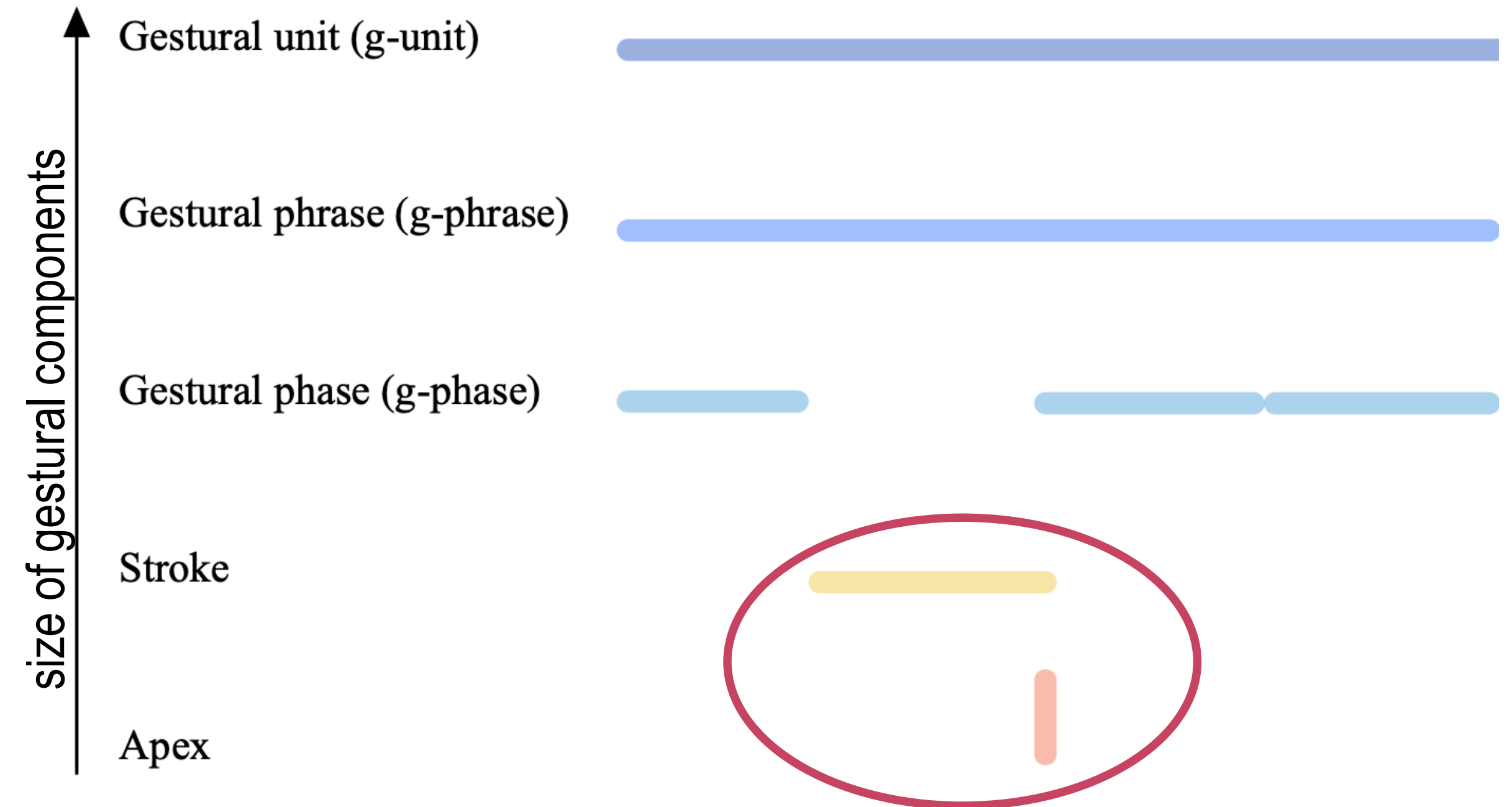


**Co-speech gestures** are „visible bodily action“  
accompanying speech (Kendon 2004)

- performed by body, **hands**, face, eyebrows

Gestures comprise of multiple **hierarchically**  
ordered components (Kendon 2004)

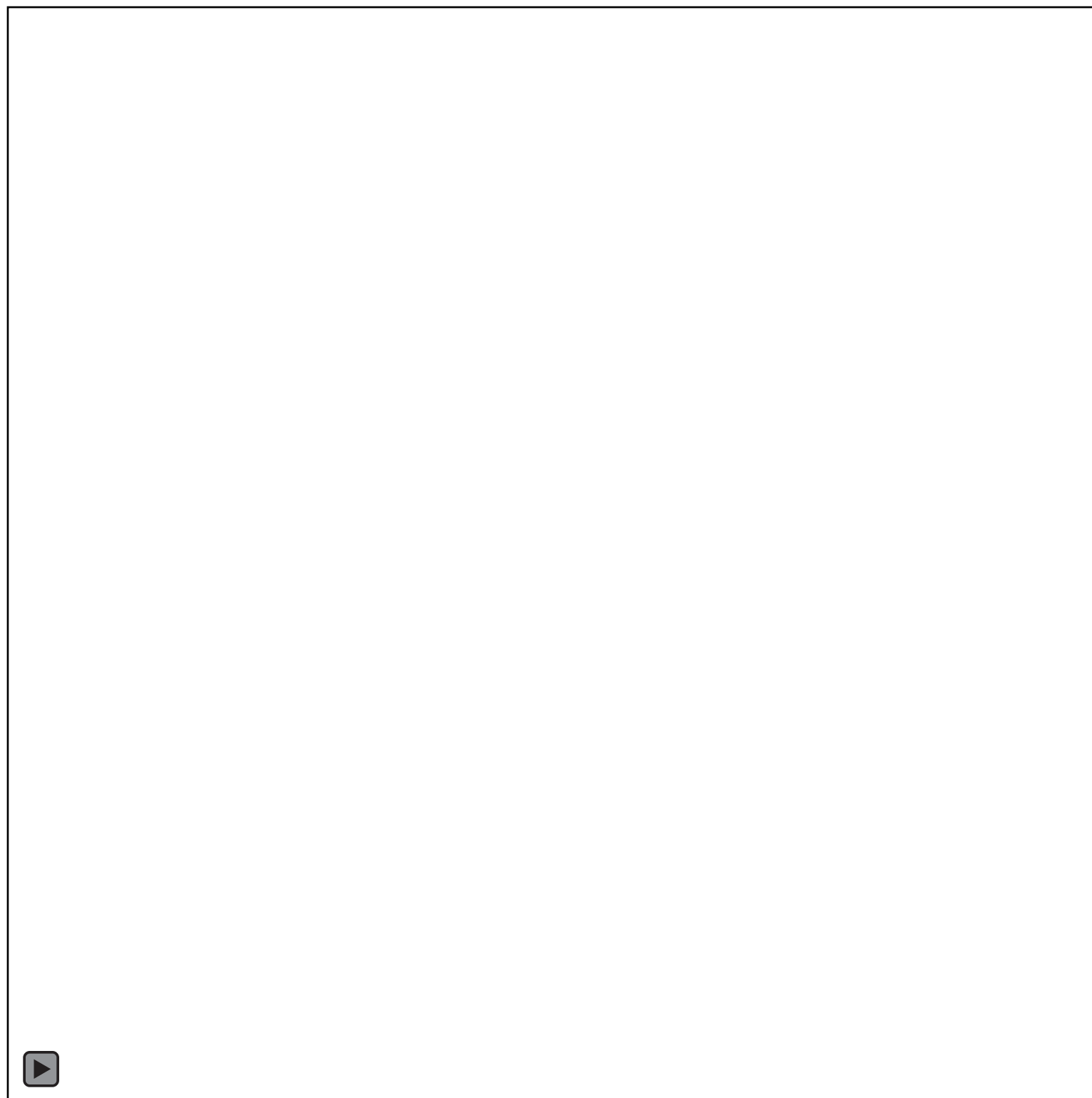
- **apex** of the stroke: gestural peak, temporal point of  
no velocity → direction change (Rohrer et al. 2020)



# Gesture types – The gesture type quartet (McNeil 1992, 2006)

## Iconic – Metaphoric – Deictic – Beat

- images of concrete entities and/or actions
- formal and structural resemblance to event or objects



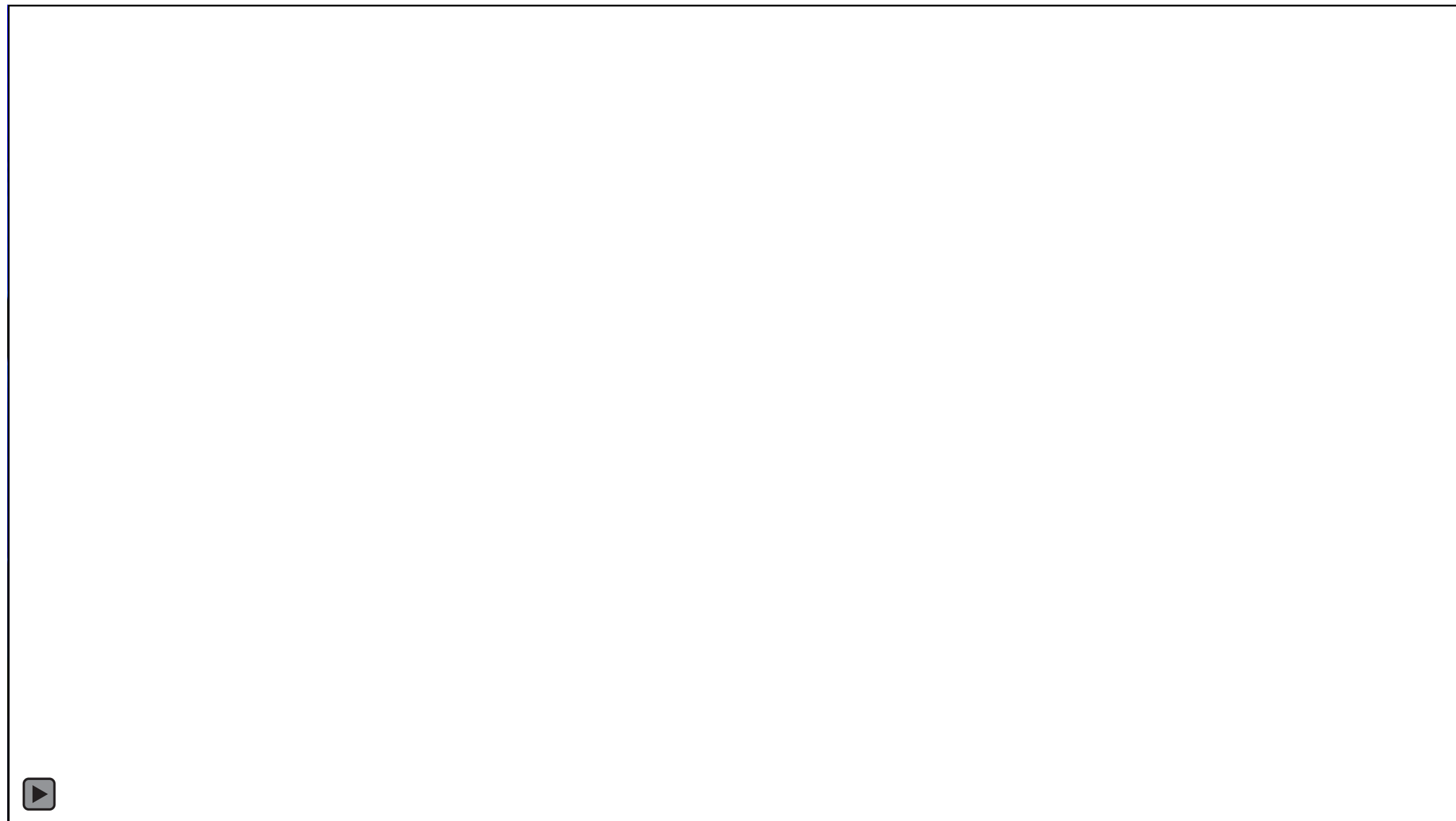
Object: a bowl, a pillar, and a smaller bowl

(SaGA corpus; Lücking et al. 2010)

# Gesture types – The gesture type quartet (McNeil 1992, 2006)

Iconic – Metaphoric – Deictic – Beat

- depictions of concrete or abstract events
- abstract meaning is presented as if it had form and/or occupied space



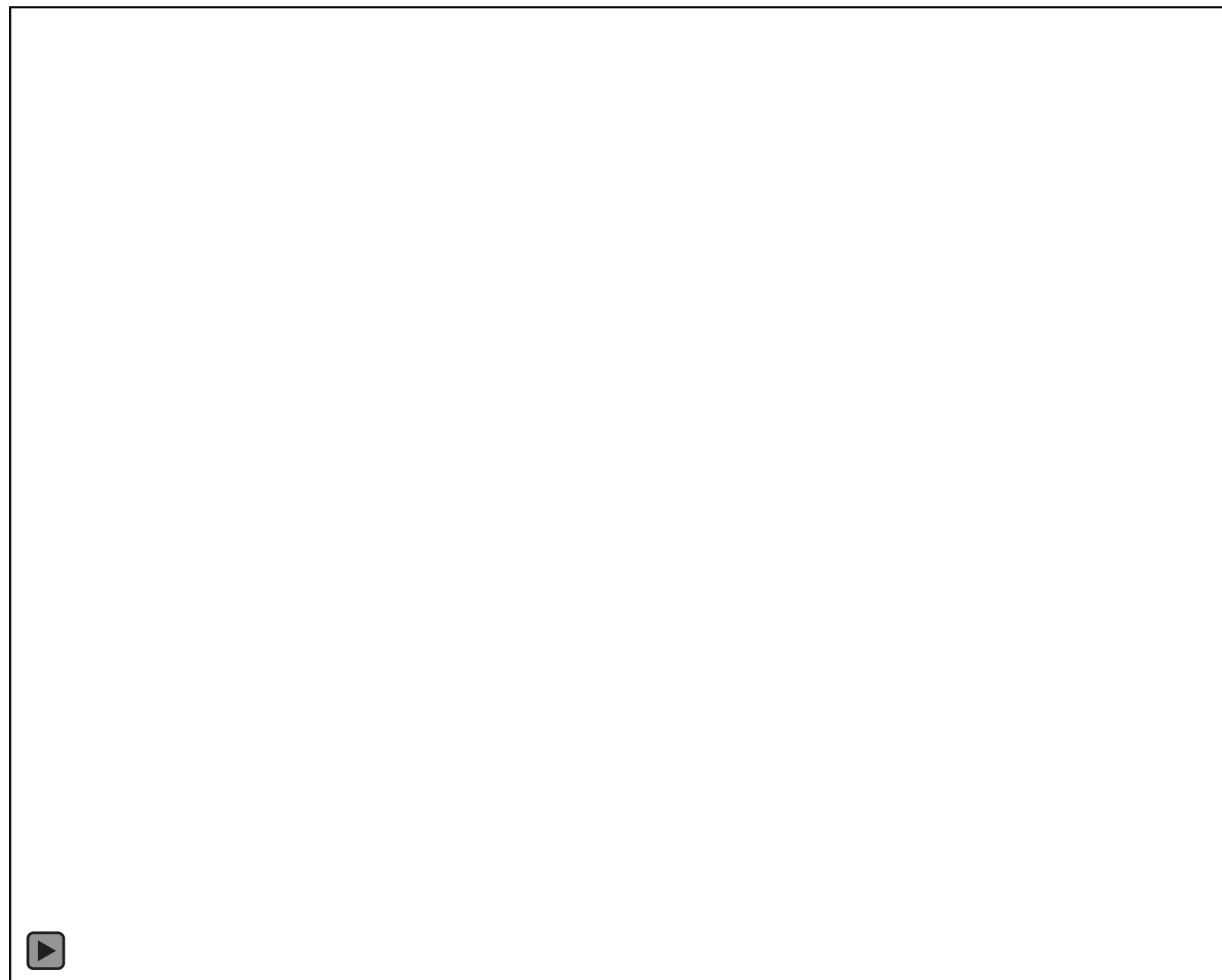
Abstract event: back (in time)

(Data from Rohrer et al. 2020)

# Gesture types – The gesture type quartet (McNeil 1992, 2006)

Iconic – Metaphoric – Deictic – Beat

- prototypically an extended ‘index’ finger
- Deixis entails locating entities and actions in space vis-à-vis a reference point



Direction: left, completely left, a little to the left

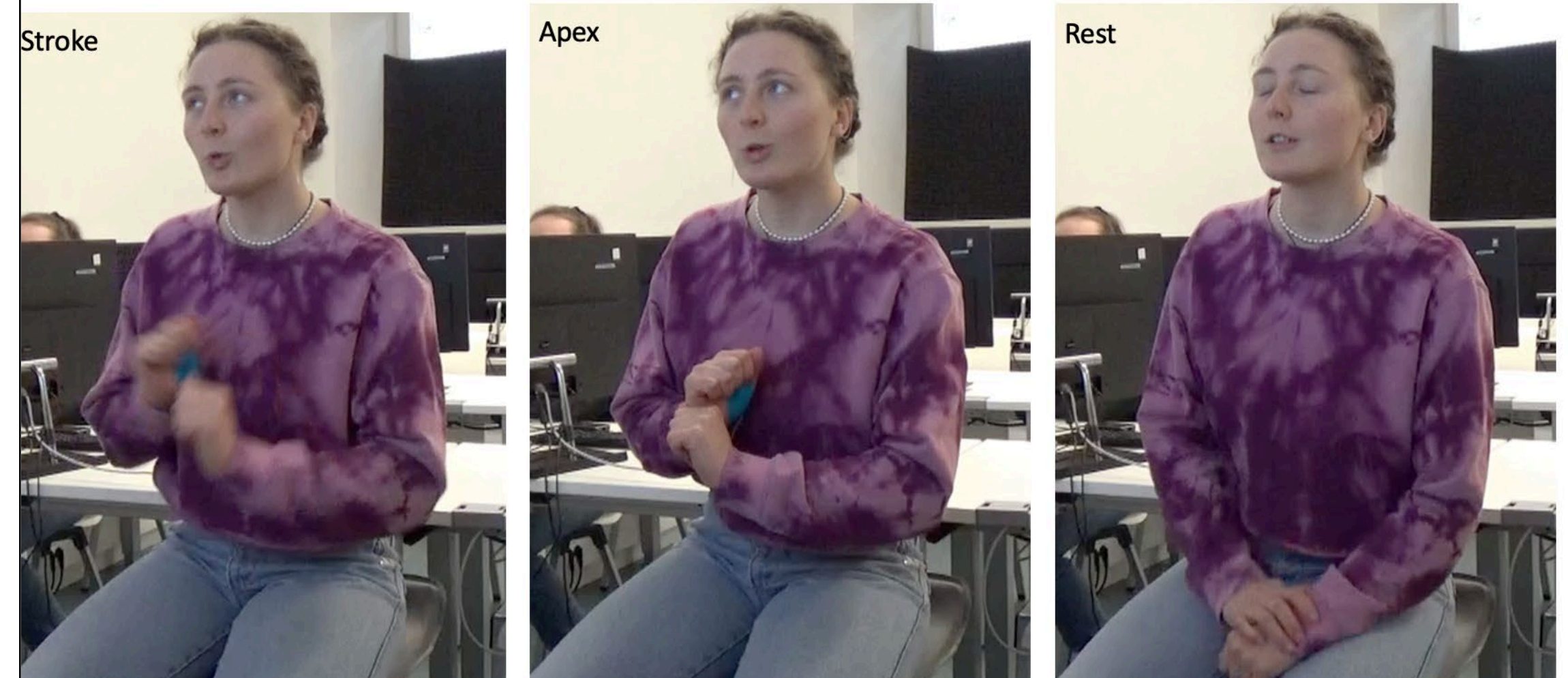
(SaGA corpus; Lücking et al. 2010)

# Gesture types – The gesture type quartet (McNeil 1992, 2006)

Iconic – Metaphoric – Deictic – Beat (non-referential)

- the hand appears to beating time
- mere flicks of the hand(s) up and down or back and forth
- zeroing in rhythmically on the prosodic peaks (purely speech-related; discourse functionality)

Beat: movement of both fists



(Data from MultIS)

# Gesture Dimensions

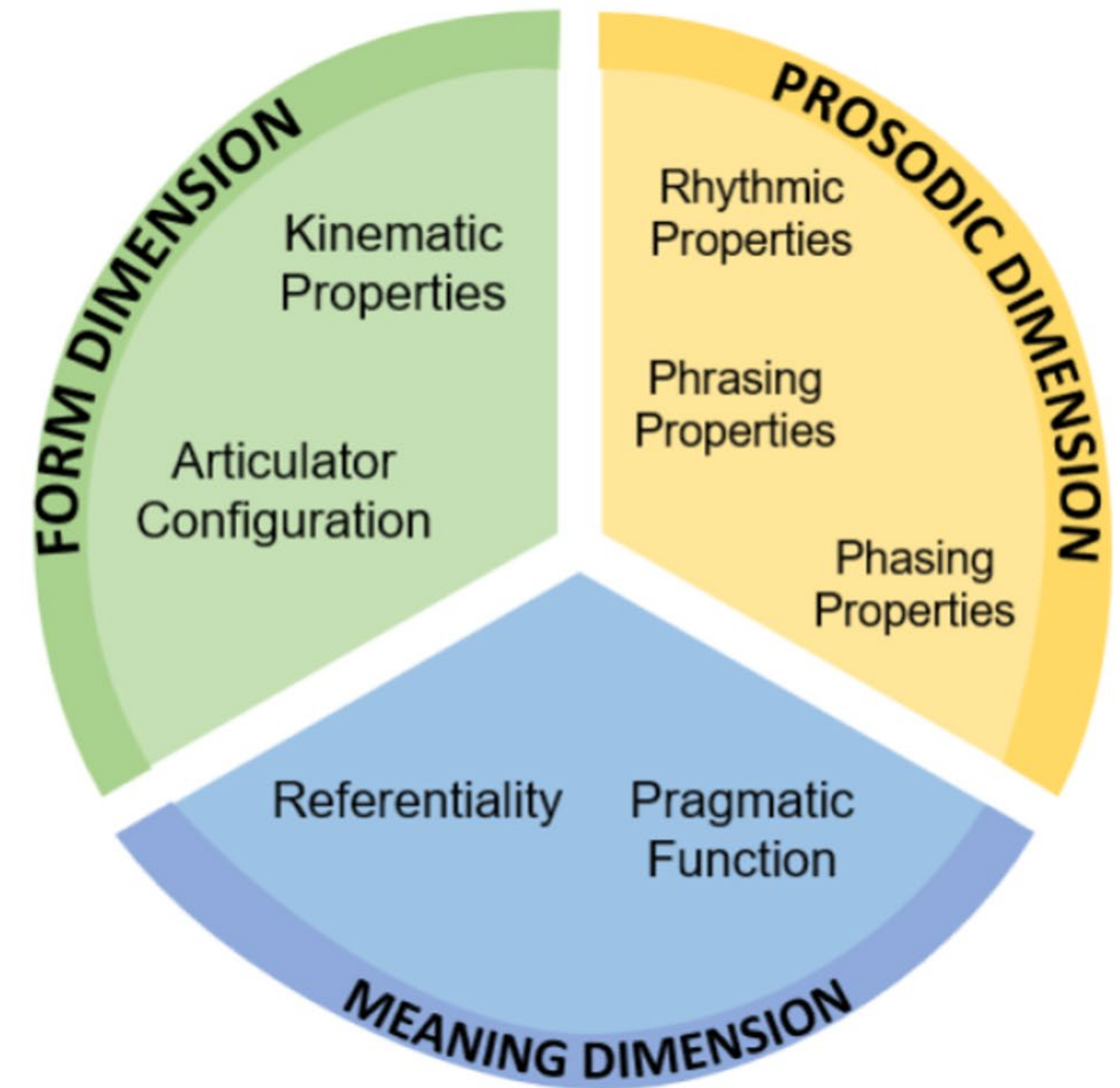
## Dimension:

“We often find iconicity, metaphoricity, deixis and other features mixing in the same gesture.”

(McNeill 2006)

## MultiModal MultiDimensional (M3D)

- New approach (Rohrer et al. 2020, Rohrer 2022)
- **Gestures do not have to be classified into one distinct type**
- Gestures have multiple dimensions
- They can receive multiple labels per dimension
- Classification is based on McNeill 1992, 2006
- Comprehensive system for labelling gestures



**Dimensions of the M3D labelling system and corresponding properties.**

OSF Project: The MultiModal MultiDimensional (M3D) labeling system  
Patrick Louis Rohrer, Ingrid Vilà-Giménez, Júlia Florit-Pons, Glenda Gurrado, Núria Esteve Gibert, PeiLin Ren, Stefanie Shattuck-Hufnagel, Pilar Prieto

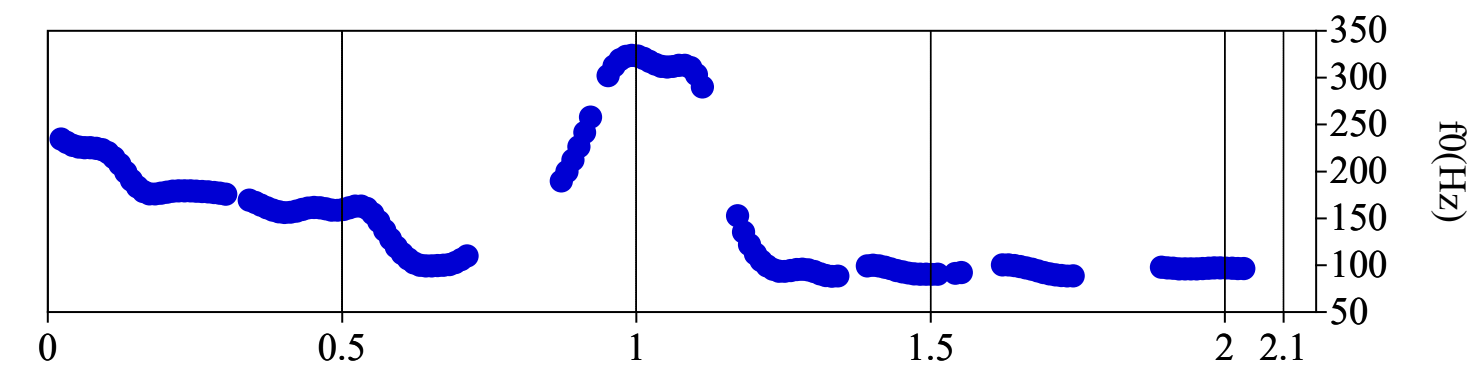


## Gesture – Prosody – Link

Gesture



Prosody



# Background - The Prosody-Gesture Link

McNeill (1992) on synchrony:

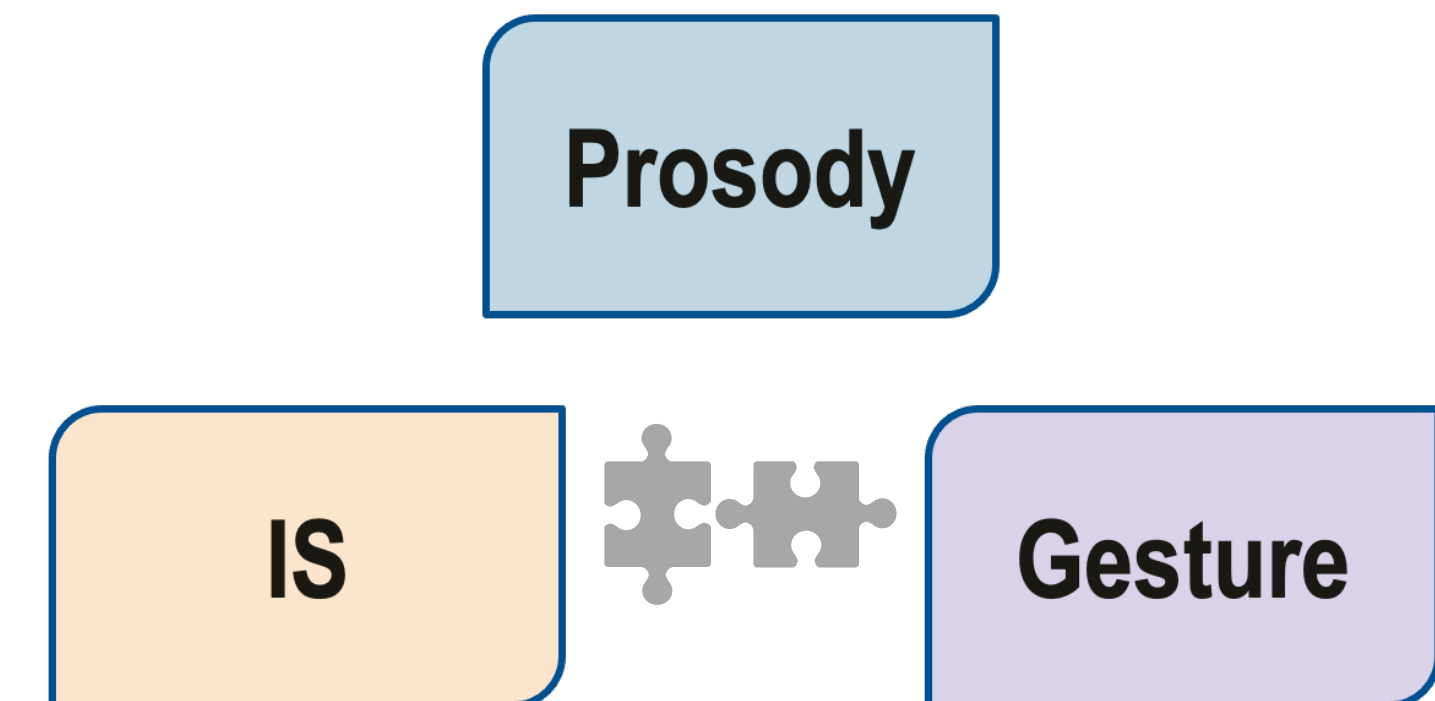
- Gestures and speech are two modalities of the same framework
- **Phonological synchrony rule**: „the stroke of the gesture precedes or ends at, but does not follow, the phonological peak syllable of speech“ (McNeill 1992, p. 26)

Gestures and speech have been empirically found to occur together

- Temporal synchronization (Shattuck-Hufnagel et al. 2007, Loehr 2012)
- Synchronization of the **stroke** and **pitch accent** (McNeill 1992), but also bigger constituents (e.g. g-phrases and phonological phrases) (Loehr 2012)

Influence of IS on their synchronization (Rohrer 2022, Im & Baumann 2020 on English)

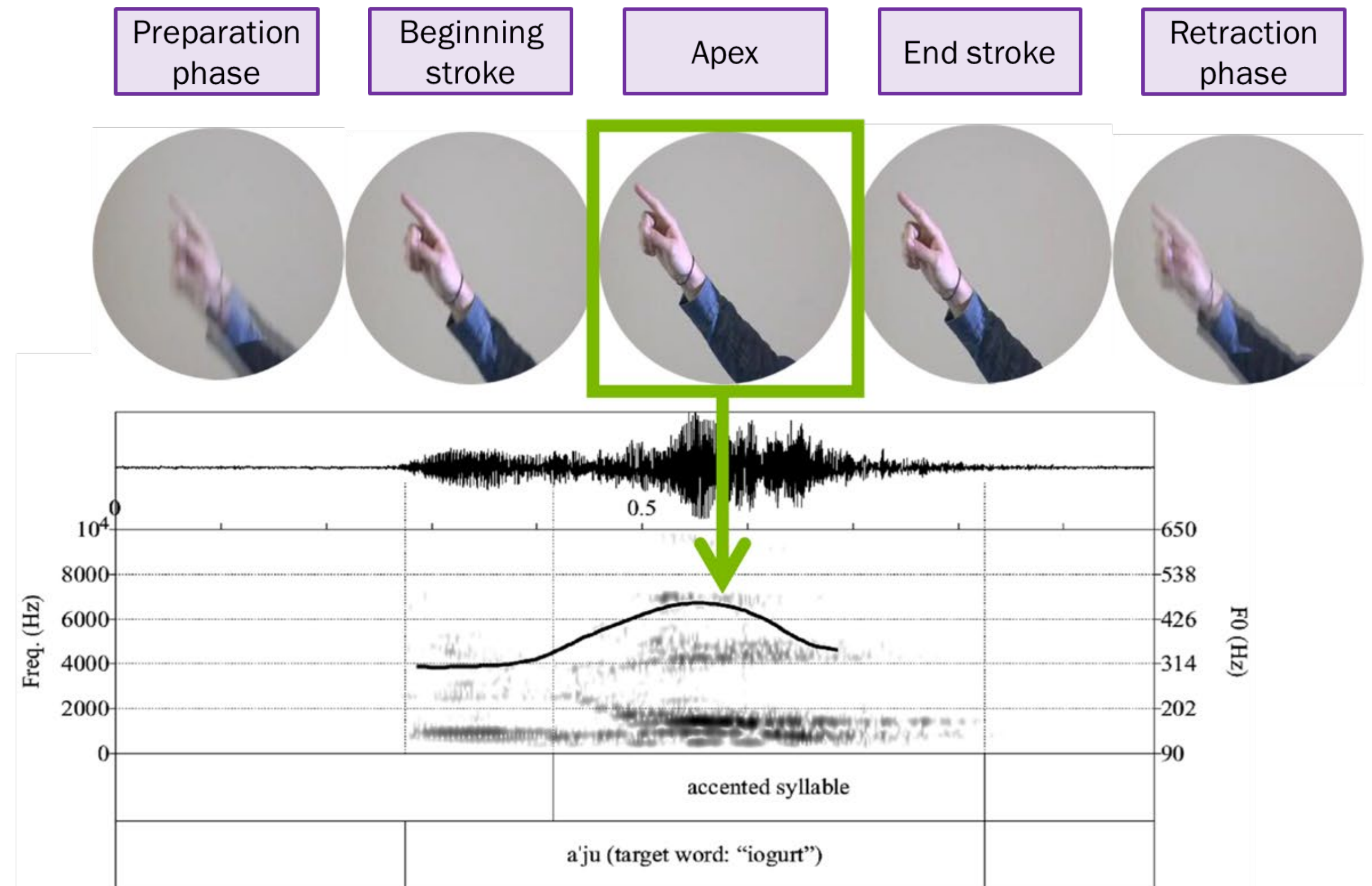
- **Newness facilitates the occurrence of gestures** (and pitch accents)
- Gestures occur less often than pitch accents



# Temporal alignment between prosodic and gesture prominence

Previous studies have shown that **prosodic** and **gesture prominence** generally are aligned in many languages (semantically, pragmatically, **phonologically**).

The gesture apex occurs at the intonation F0 peak (Esteve-Gibert & Prieto, 2013): Phonological Synchrony Rule (McNeill, 1992)

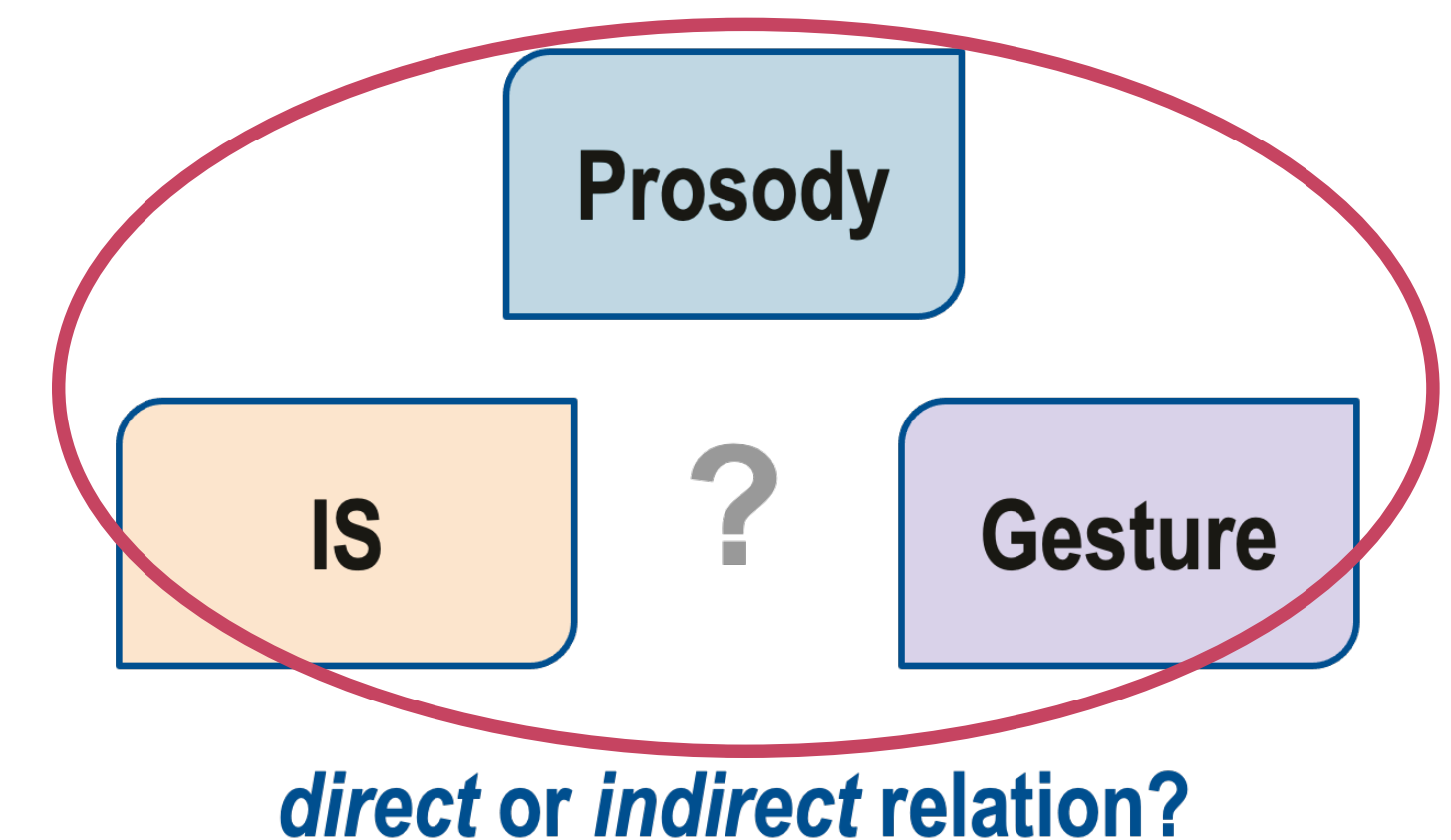
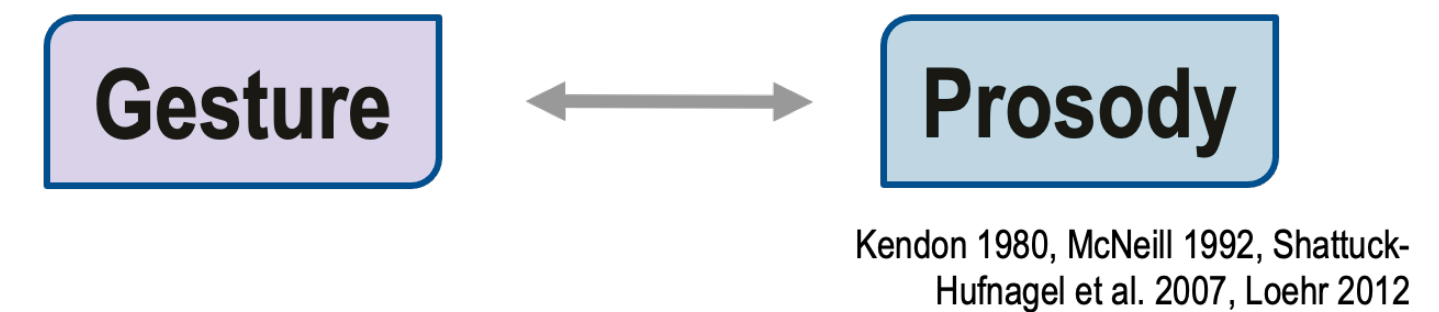
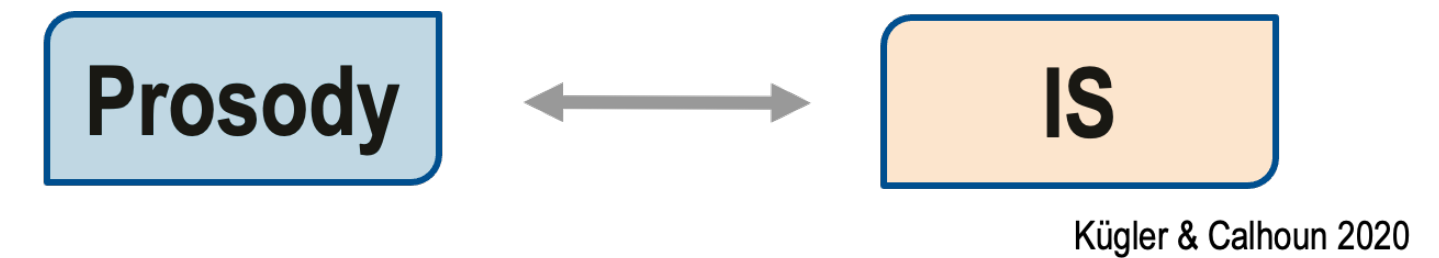


(Apex-accent alignment taken from Esteve-Gibert & Prieto 2013)

# Research Questions

How is prominence (focus) encoded multimodally?

- Q1: Does prominence influence the occurrence of gestures in spontaneous German speech?
- Q2: Is pitch accentuation temporally aligned with gesture apexes in German and does prominence influence this alignment?
- Q3: Are different types of gestures (iconic and non-referential) affected by prominence in the same way and extent?
- Q4: Are different degrees of prominence encoded with different degrees multimodal prominence?



## A corpus study

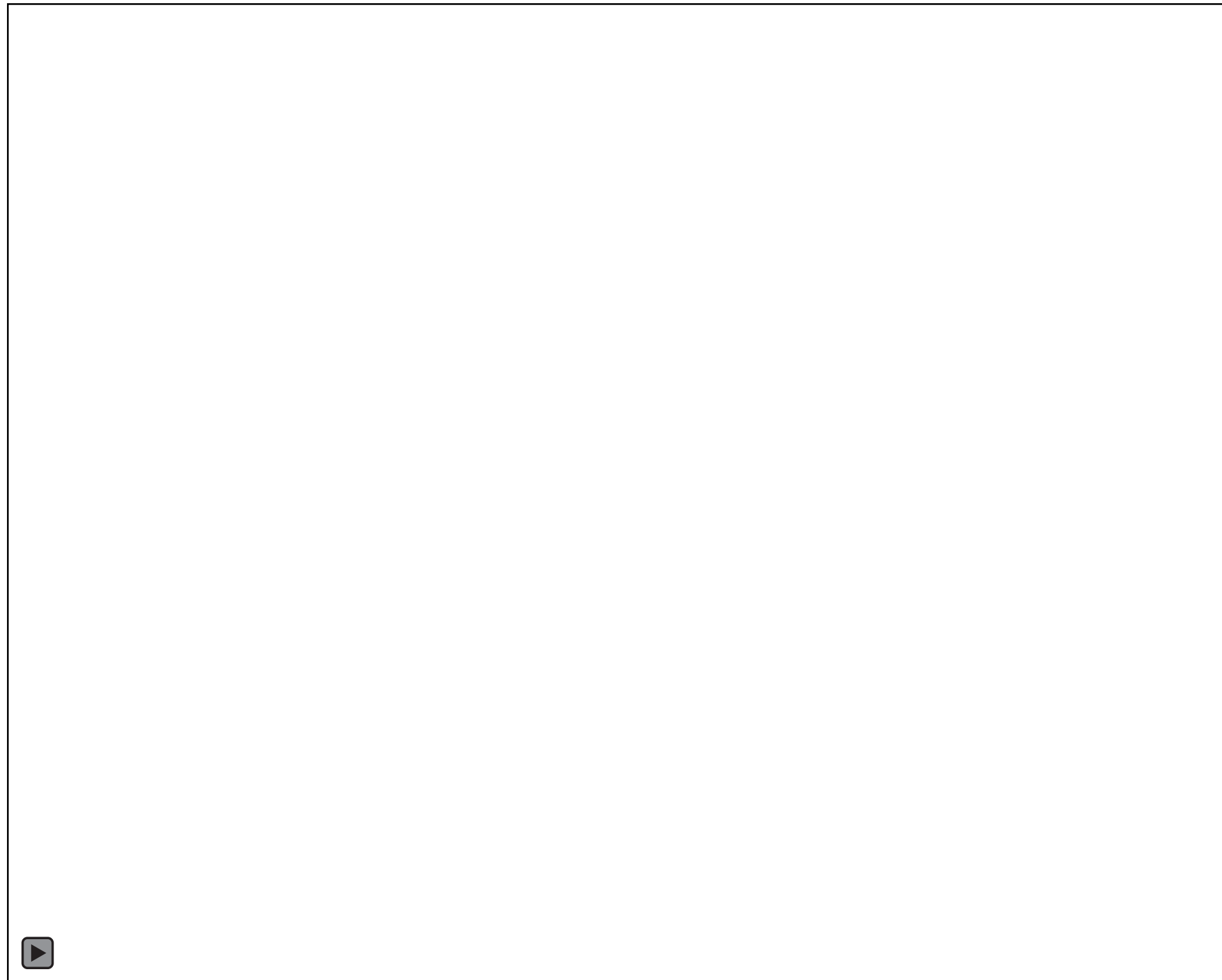
Joint work with Alina Gregori  
(Gregori 2022; Kügler & Gregori to appear)

**Alina Gregori**



# The Corpus

VR environment: Bus drive



Conversation: Route description



Bielefeld Speech and Gesture Alignment (**SaGA**) corpus (Lücking et al. 2010)

- audio-visual corpus containing German spontaneous speech conversations
- **VR town environment** to provide a stimulus and direction-giving task
- 204 min of dialogues, gesture types provided

# Method

Conversations were annotated for:



## IS Annotation

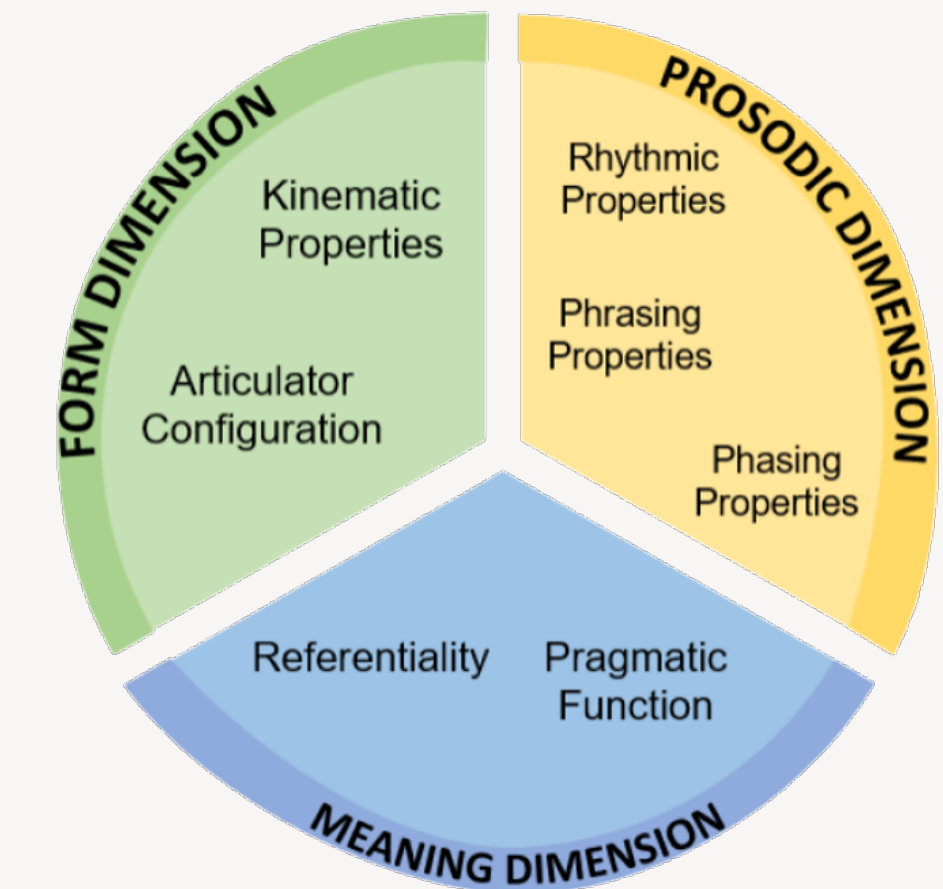
IS levels annotated in accordance with  
Götze et al. (2007) and Krifka (2008)

## Prosodic Annotation

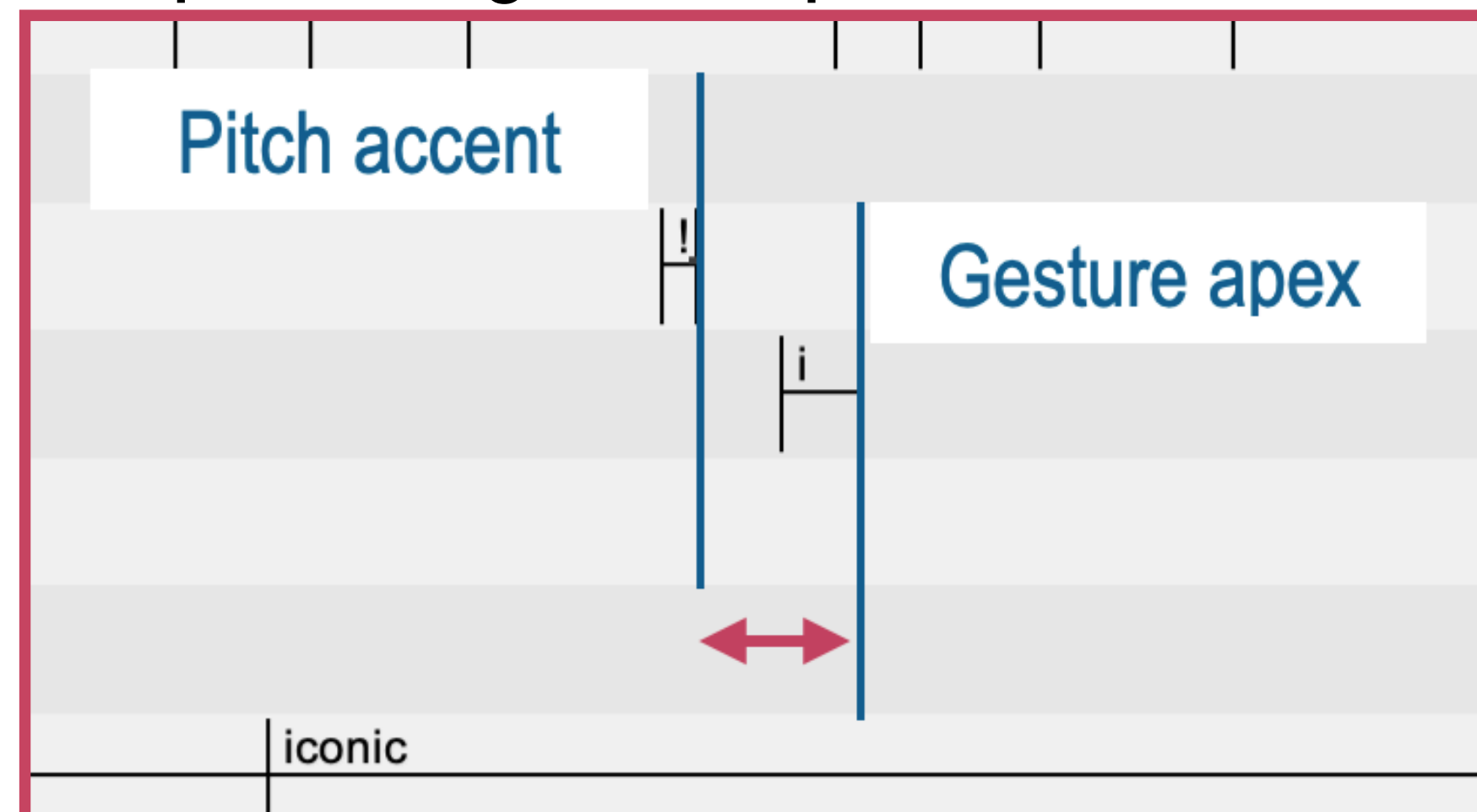
Phonological annotation in GToBI (Grice et al. 2005)  
Phonetic analyses

## Gesture Annotation

The MultiModal  
Multidimensional (M3D)  
labelling System  
(Rohrer et al. 2020)



Temporal alignment procedure:



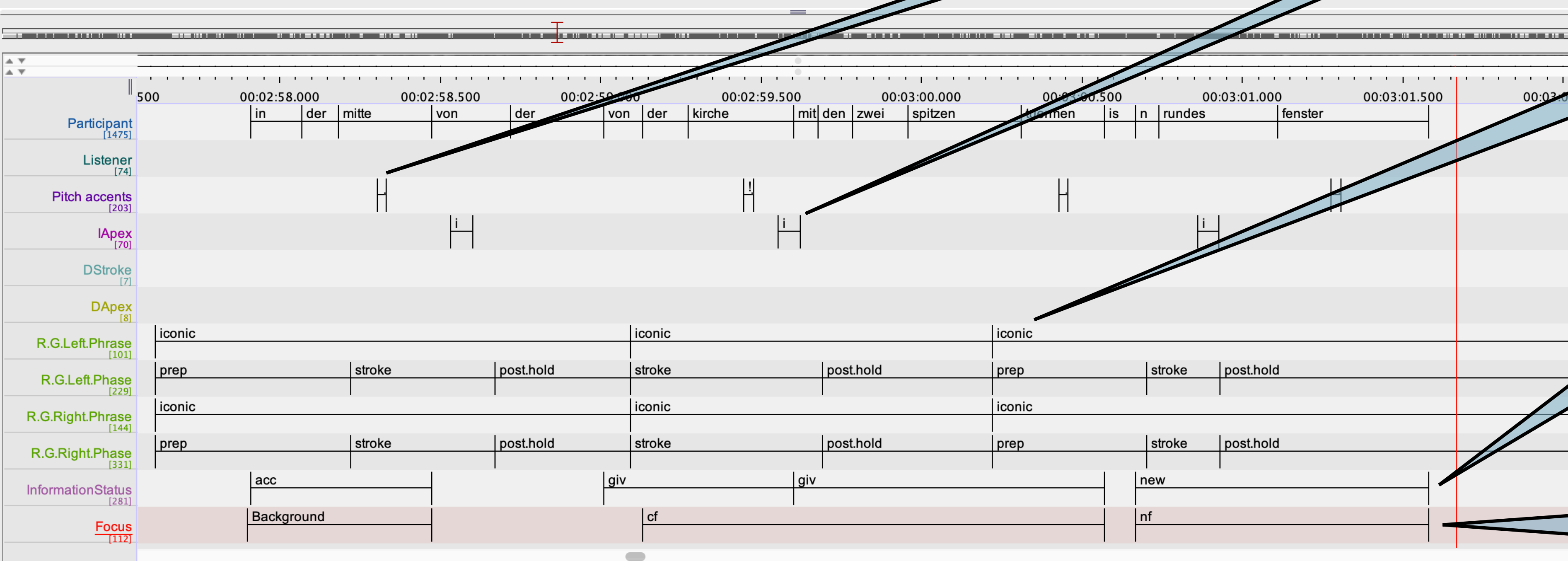
Analysis of iconic (referential) and  
beat (non-referential) gestures

# Corpus study – Hard Numbers



geht · dann · genau · es · geht · dann · irgendwann · nicht · mehr · geradeaus · weiter · am · fluss · lang · sondern · die · strasse · knickt · dann · nach · rechts · ab · der · folgst · du · dann · einfach · nach · rechts · haelst · dich · aber · immer · noch · soweit · links · wie · es · geht · also · hier · is · der · fluss · daneben · die · strasse · die · geht · irgendwann · nach · rechts · hier · komm · n · paar · haeuser · sind · ueberall · haeuser · und · dann · haelst · du · dich · aber · links · und · wenn · du · dann · dich · links · umdrehst · kommst · du · auf · so · n · grossen · platz · und · aeh · auf · der · linken · seite · von · dem · platz · steht · eine · kirche · auf · der · rechten · seite · steht · auch · ne · kirche · die · linke · sieht · son · kleines · bisschen · aus · wie · ne · moschee · naja · die · hat · in · der · mitte · ein · turm · der · oben · n · rundes · dach · hat · mit · nem · kreuz · drauf · aber · und · auf · der · rechten · seite · is · ne · andere · kirche · die · hat · zwei · kirchtuerme · die · spitz · sind · auf · dem · rechten · der · beiden · kirchtuerme · is · auch · n · kreuz · auf · dem · linken · nich · wolln · die · das · gleich · noch · wissen · in · der · mitte · von · der · von · der · kirche · mit · den · zwei · spitzen · tuermen · is · n · rundes · fenster · n · rundes · kirchenfenster · gut · das · is · der · platz · mit · den · zwei · kirchen · danach · aehm · kannste · von · dort · aus · einfach · weiter · geradeaus · laufen · ueber · den · platz · drueber · und · kommst · dann · ja · in · strassenfolgen · ja · lass · mich · nur · kurz · nachdenken · aeh · mmm · tja · ich · glaube · du · bist · dann · wenn · du · weitergelaufen · bist · bist · du · auf · so · nen · t · foermige · kreuzung · gekommen · musst · du · dich · links · drehen · und · ahem · wenn · du · dann · weiter · geradeaus · gehst · kommt · auf · deiner · rechten · seite · n · grosser · park · das · siehst · da · siehst · du · erstmal · nur · ne · hecke · auf · der · rechten · seite · dann · laeufst · du · weiter · bis · du · irgendwann · auf · der · rechten · seite · son · eingangstor · zu · dem · park · findest · dort · drehst · du · dich · rein · laeufst · durch · den · park · durch · du · kannst · im · prinzip · rumlaufen · wie · d · willst · inner · mitte · isn · grosser · see · und · aehm · um · den · see · fuehrt · links · n · weg · rum · rechts · n · weg · weg · rum · so · im · kreis · und · aehm · wenn · du · also · und · durch · das · tor · reingegangen · bist · in · dem · park · drin · bist · auf · den · see ·

00:03:01.667 Selection: 00:00:00.000 – 00:00:00.000 0



Pitch accent type  
(n = 4394)

Apex (n = 2402)

Gesture type (n = 2402):  
Iconic (n = 1627)  
non-referential (n = 775)  
other types

Info status (n = 3939):  
new  
accessible  
given  
NoInfo (n = 1085)

Focus (n = 2773):  
cf = contrastive focus  
nf = new-information focus  
Background (n = 2251)



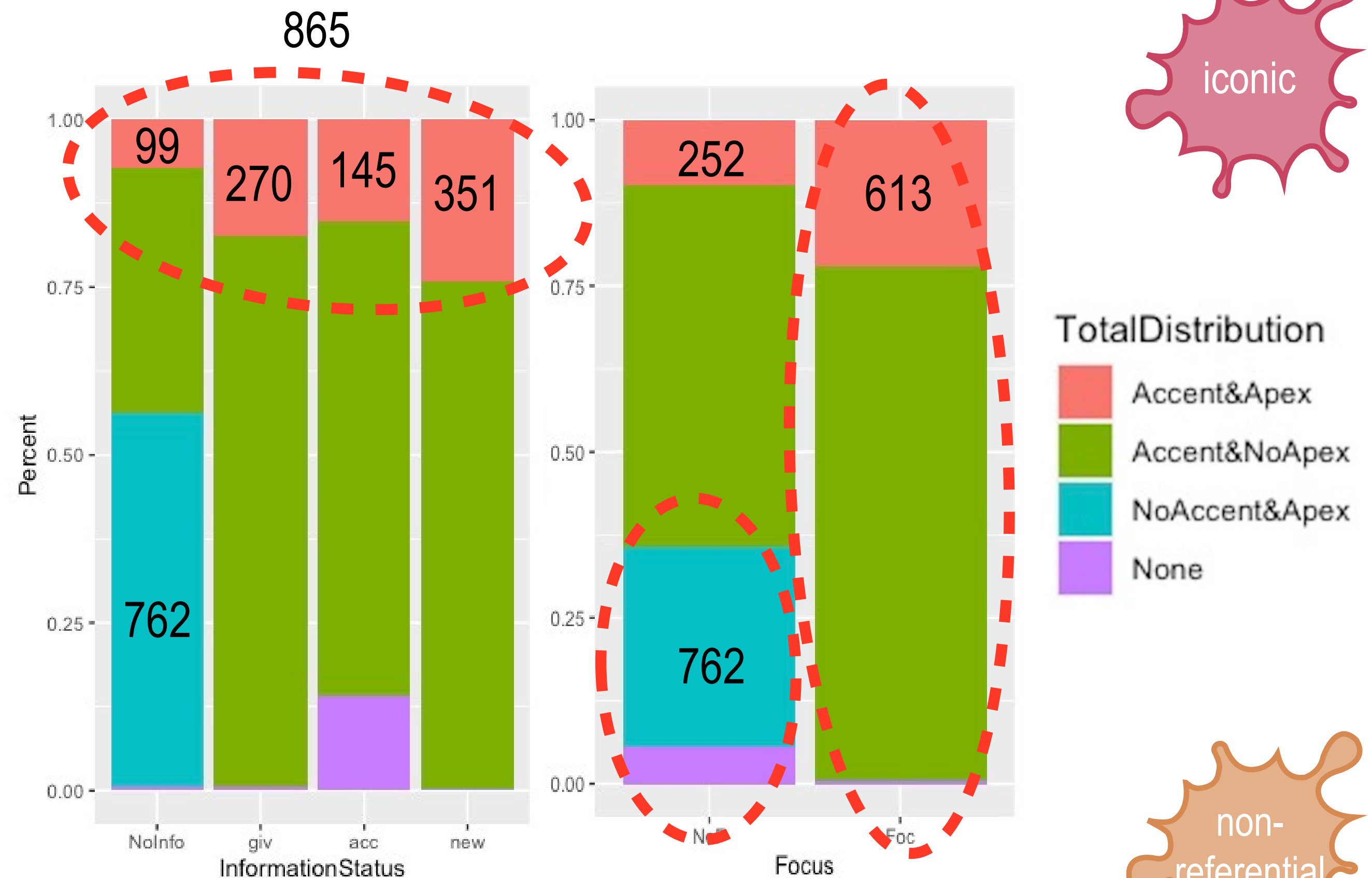
# Results – Distribution of Gestures

## Gesture occurrence split by **Information status**

- *Given* and *New* always marked prosodically
- Pragmatically more prominent: more Accent + Apex
- Preferable pitch accents / gestures alone
- In *NoInfo* gestures predominantly without accent (55%)

## Gesture occurrence split by **Focus**

- *Focus* always marked prosodically
- In 25% additionally by gesture
- Preferably pitch accents / gestures alone
- In *non-focus* gestures without accent (33%)



Similar picture for both IS levels; lower percentage

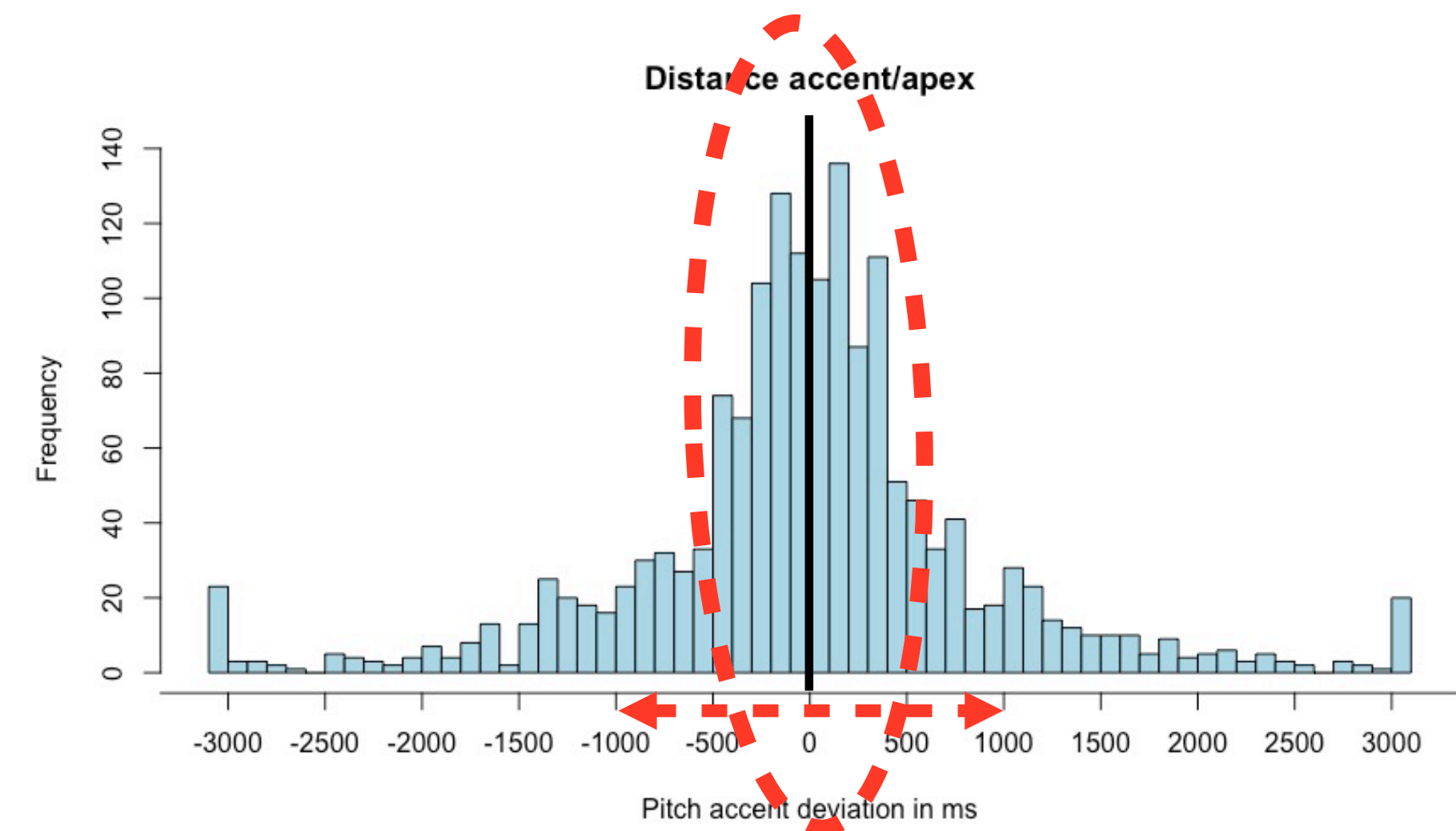
→ GREGORI 2022

# Results - Temporal Alignment

Pitch accents and apexes tend to occur near each other

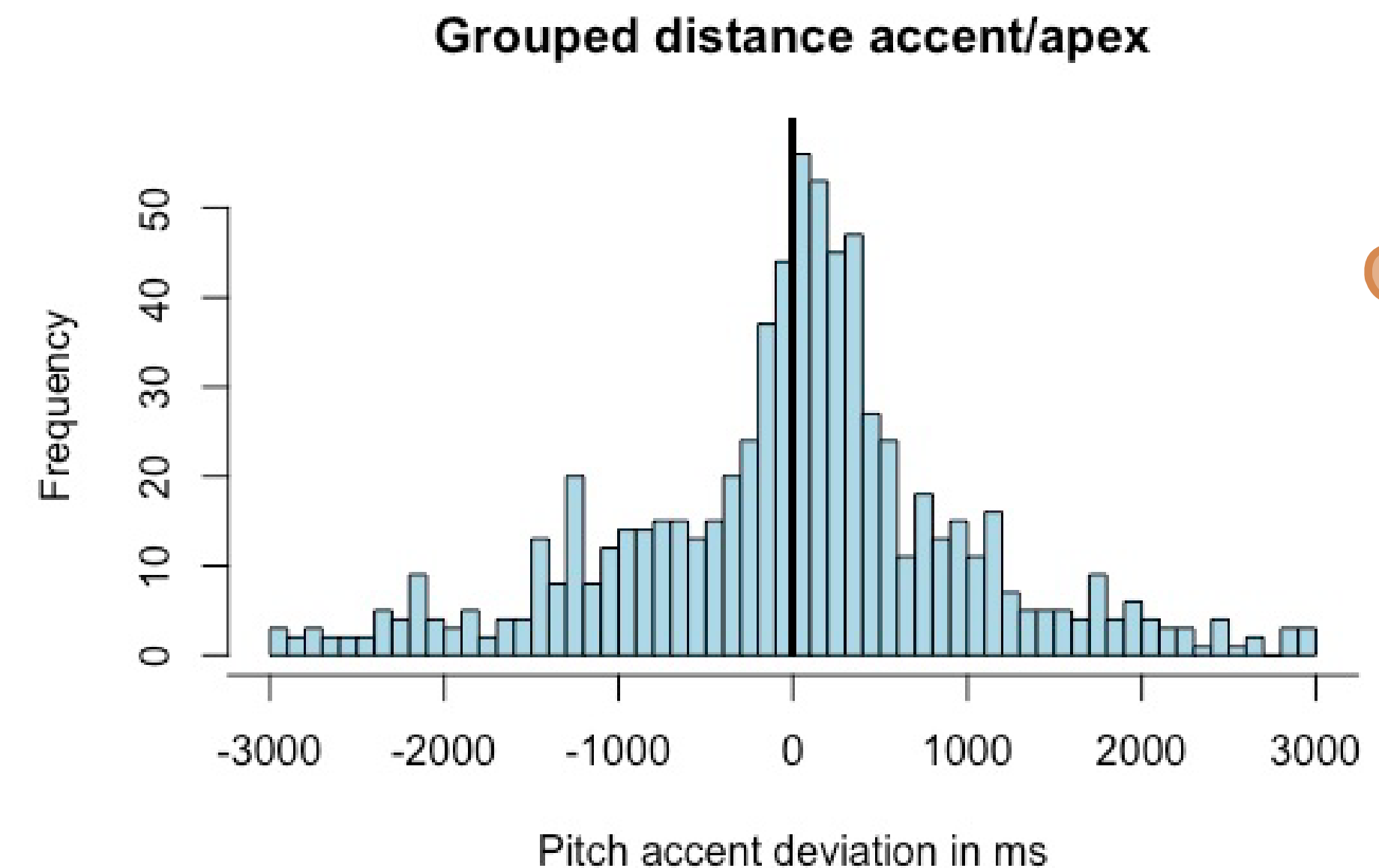
## Iconic

- mean deviation 7ms, pitch accent before apex
- standard deviation 375ms
- 50,3% PA after apex; 0,1% exactly aligned
- 78,4% within one second distance



## Non-referential

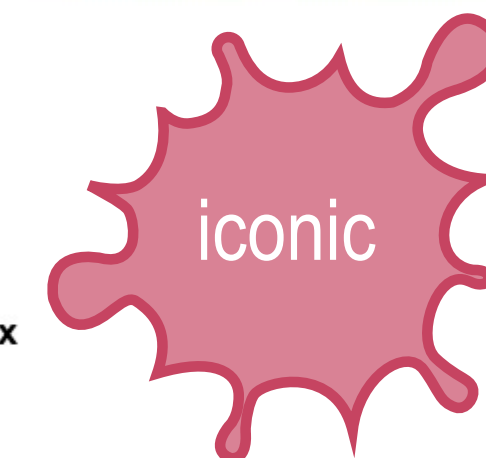
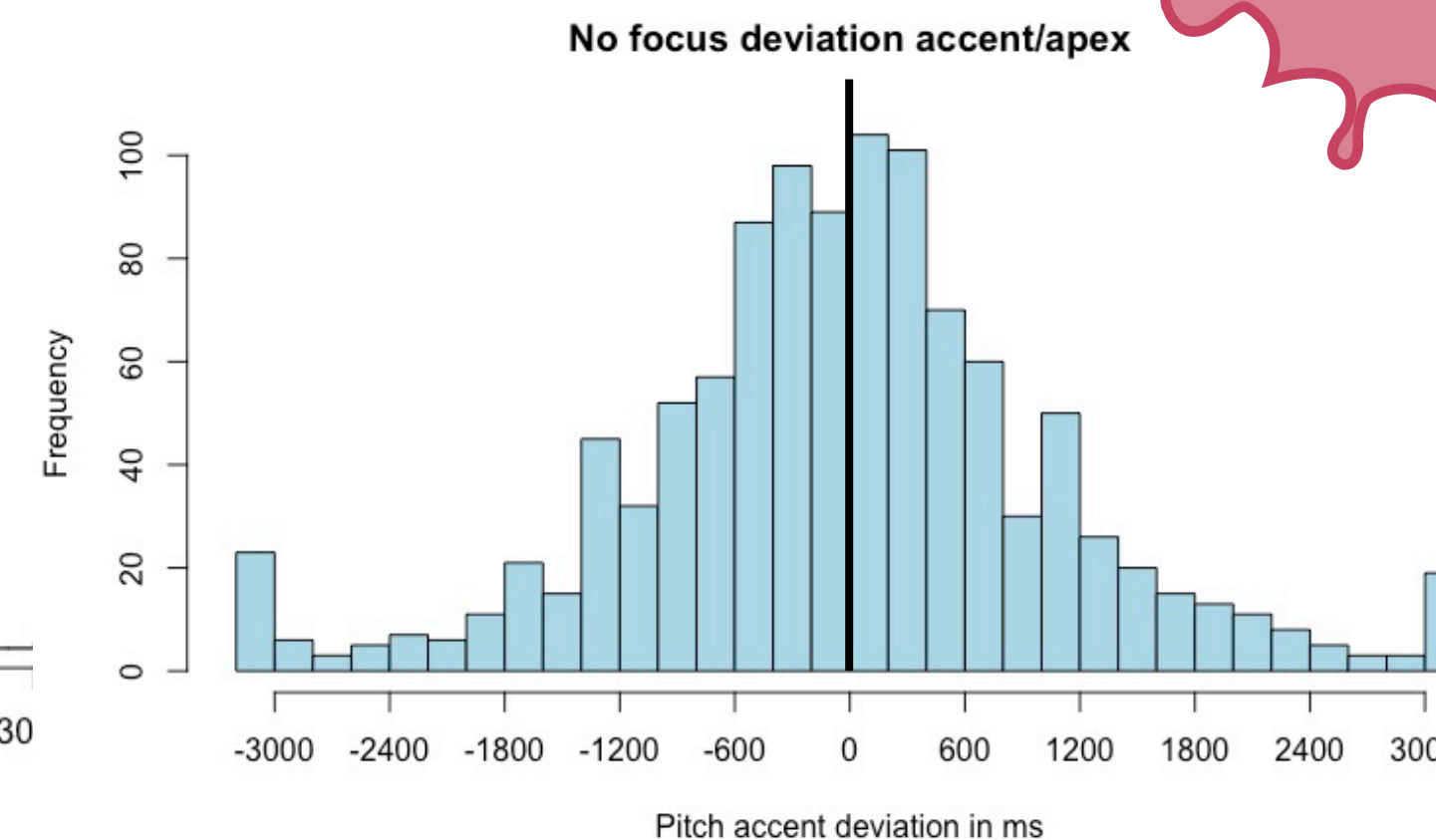
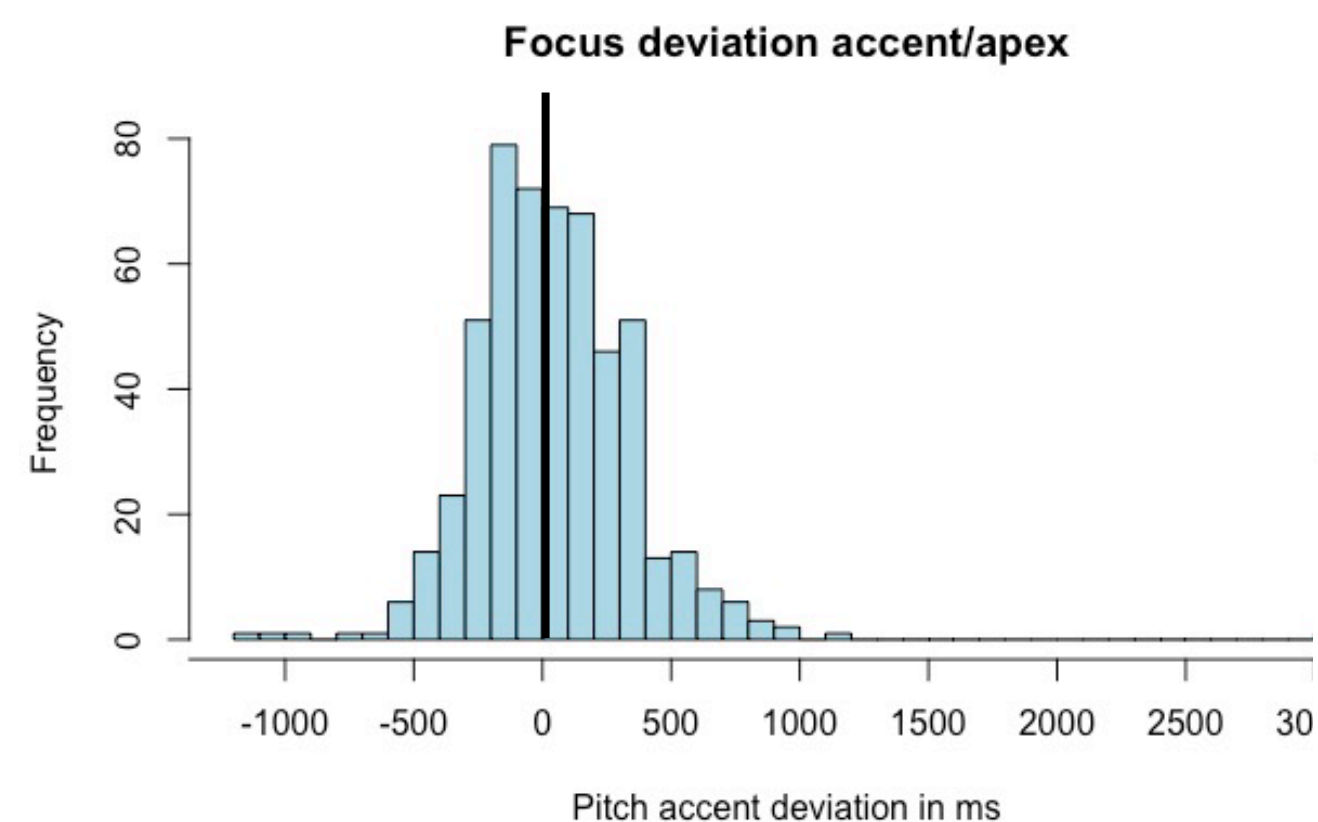
- mean deviation 38ms, pitch accent after apex
- standard deviation 385ms
- 54,5% PA after apex; 3,6% exactly aligned
- 66,8% within one second distance



# Results - Temporal Alignment

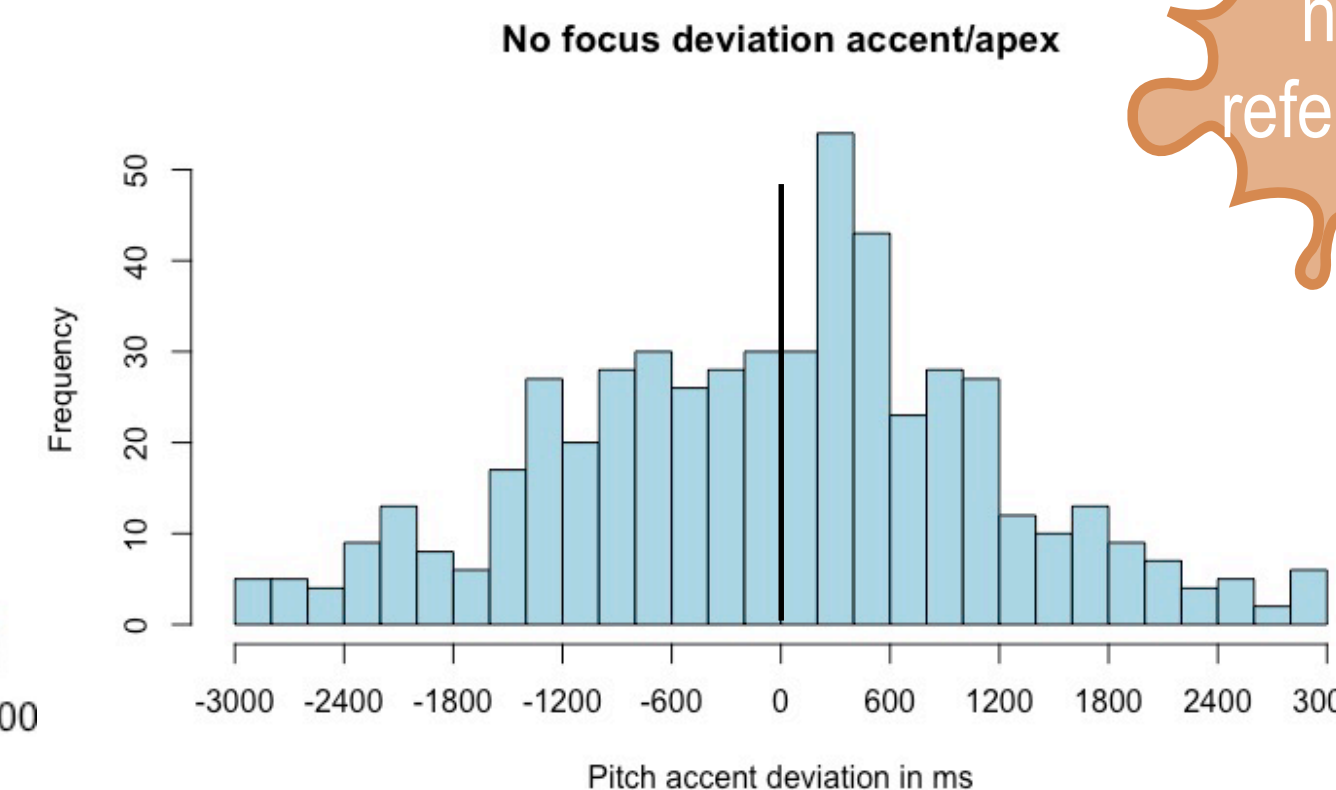
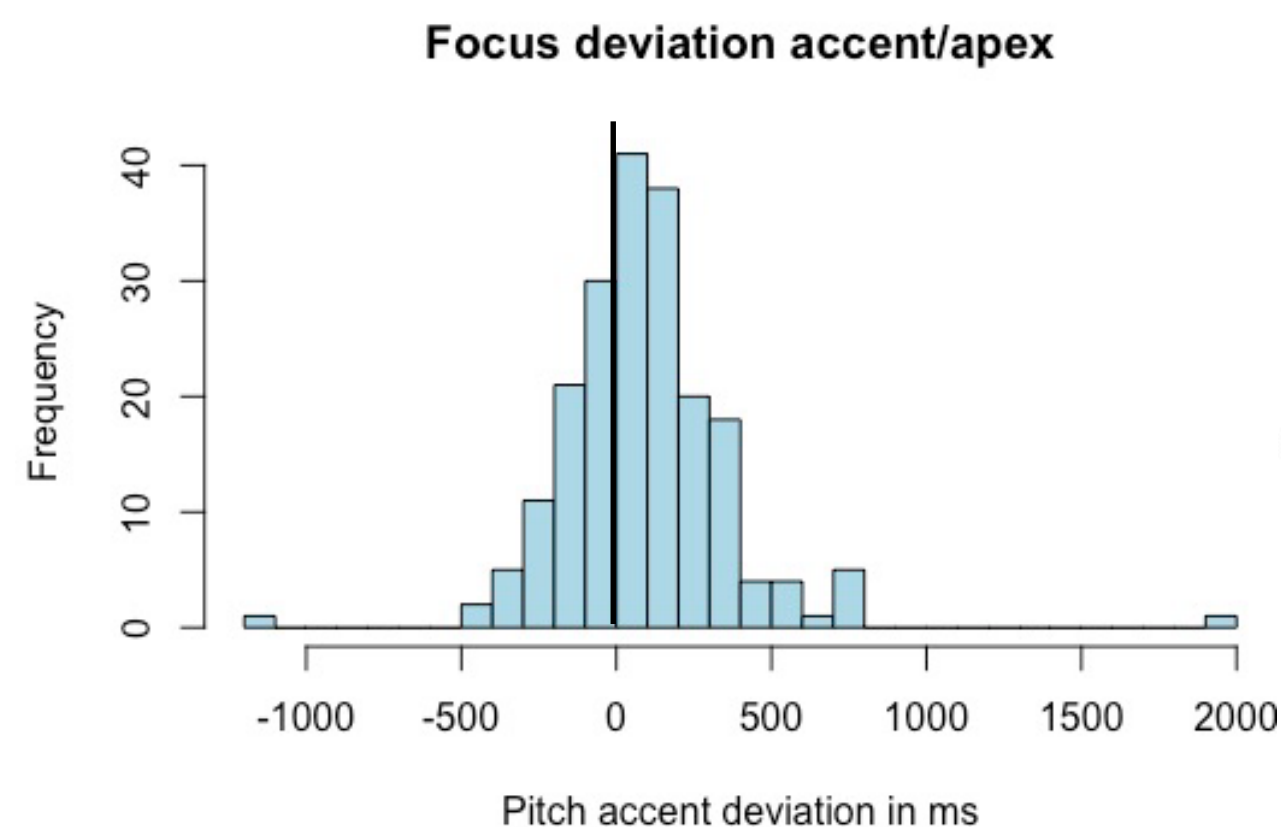
## Focus - Iconic

	Mean	Standard deviation	PA after apex	within -500 & 500ms
Focus	41ms	267ms	53 %	91,3 %
Non-focus	-9ms	416ms	49,1 %	44,7 %



## Focus - Non-referential

	Mean	Standard deviation	PA after apex	within -500 & 500ms
Focus	81ms	220ms	65 %	95 %
Non-focus	22ms	432ms	51,6 %	35 %



# Corpus study: Discussion

Pitch accents and apexes tend to occur close to each other

in line with [Loehr \(2012\)](#) and the phonological synchrony rule [McNeill \(1992\)](#)

Even closer when pragmatically prominent (focused)!

Similar behavior of iconic and non-referential (beat) gestures towards focus marking

→ Evidence towards M3D approach [Rohrer et al. 2020](#), [Rohrer 2022](#): a gesture has multiple dimensions contributing to the discourse structurally and semantically (see also [McNeill 2006](#)).

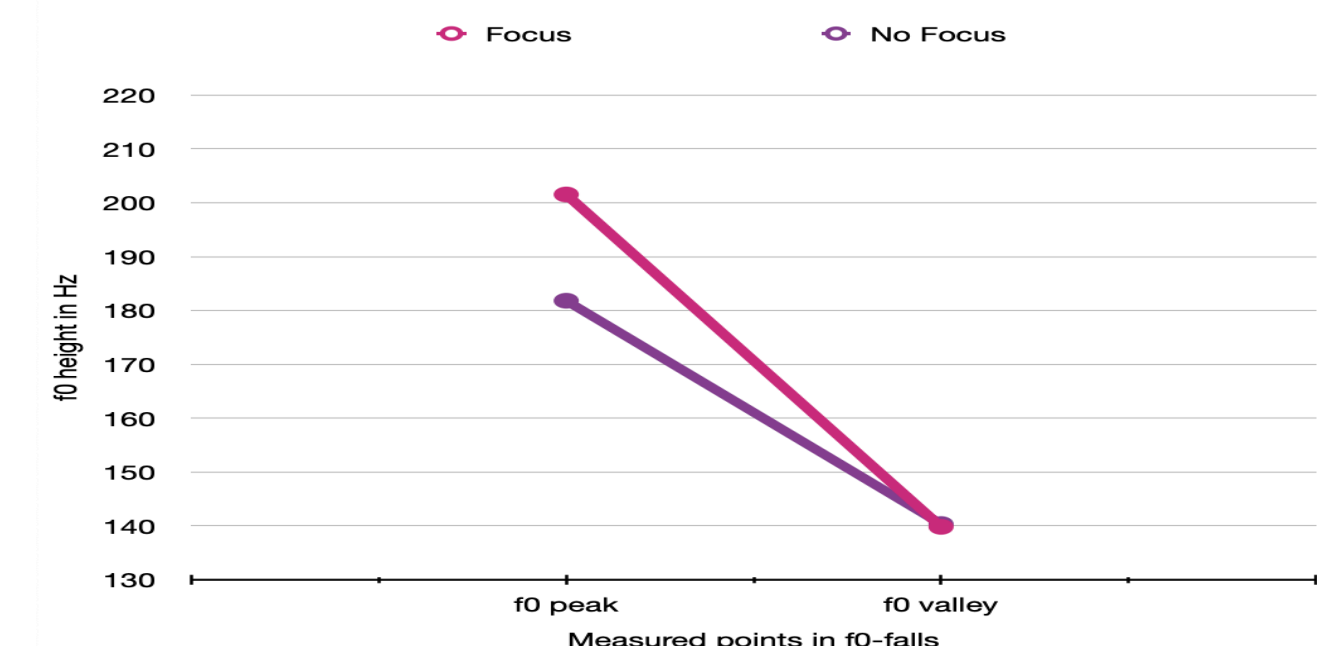
## Prominence multimodally

Multimodal Hyperarticulation ([Kügler & Gregori to appear](#)), parallel to phonetic hyperarticulation ([Lindblom 1990](#), [Hanssen et al. 2008](#)).

Acoustic and visual prominence marking cumulative instead of trading relation ([Ambrazaitis & House 2022](#)).

What about prominence degrees?

- 141 falling accents (H\*L): 100 focus, 41 non-focus
- F0-max on average 40 Hz higher in focus (cf. [Féry & Kügler 2008](#))
- Steeper slopes in focus (cf. [Hanssen et al. 2008](#) for Dutch)
- More precise f0 on focused word
- Effect of hyperarticulation (see [Lindblom 1990](#))





## An interactive production study (work in progress)

Joint work within MultIS project – Universitat Pompeu Fabra, Barcelona & Goethe University Frankfurt  
(Gregori, Sánchez-Ramón, Prieto & Kügler 2023)

**Alina Gregori**



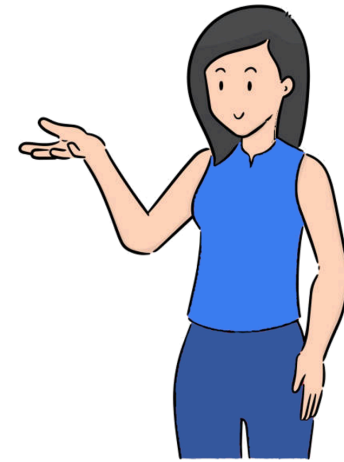
**Prof. Dr. Pilar Prieto**

**Paula G. Sánchez-Ramón**



# Method – Data elicitation

Wenn du ihren Beutel siehst, sag Maria welchen Gegenstand sie nehmen muss, um die Zeitung zu lesen.



Participants sitting on a high chair

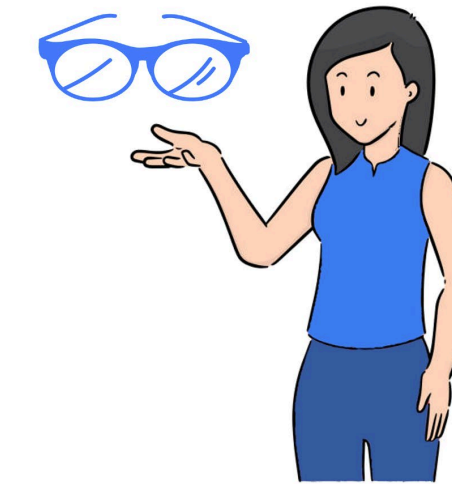
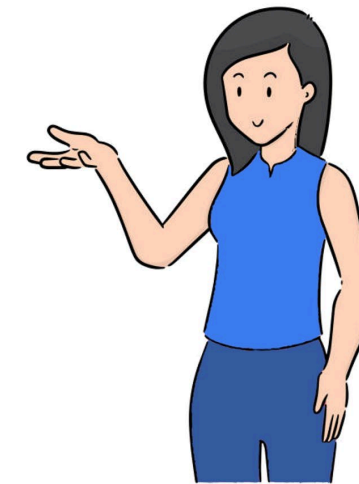
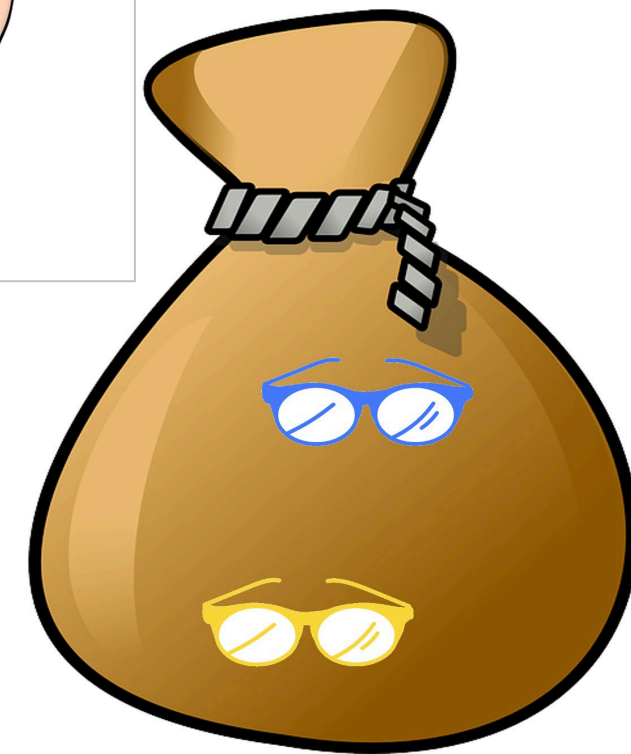
Interactive task: Maria, learning German, needs to learn instructions and colours

Context given through an action (reading newspaper)

Objects hidden in a bag

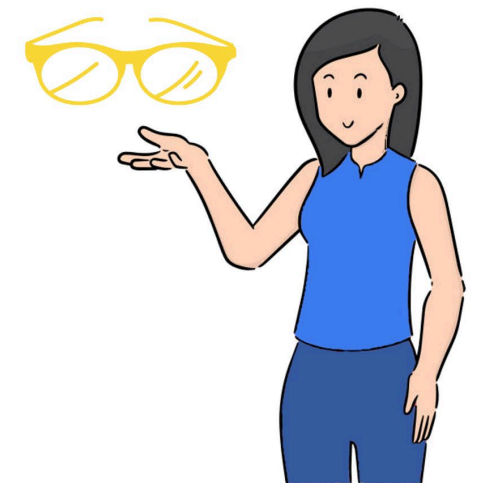
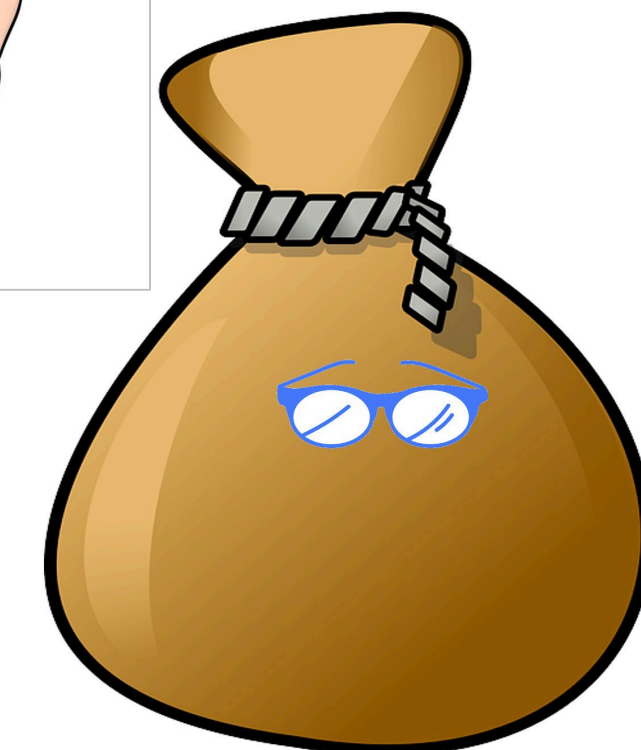
Instruction: "Maria, take the yellow glasses!"

(contrastive focus)



Task failed, Maria needs new instruction.

(corrective focus)



Task succeeded

Recordings so far:

Three speakers x

7 items x 4 conditions = 84 sentences

# Method – Data

## Elicitation of different degrees of prominence through context manipulation

### Degrees of prominence:

Background:	<i>Wofür braucht Maria die gelbe Brille?</i> Maria nimm die <u>gelbe Brille</u> [zum Zeitunglesen] <sub>foc</sub>
Narrow focus:	<i>Was braucht Maria zum Zeitunglesen?</i> Maria nimm die [gelbe Brille] <sub>foc</sub> zum Zeitunglesen
Contrastive focus:	<i>Wähle zwischen der blauen und der gelben Brille.</i> Maria nimm die [gelbe] <sub>foc</sub> Brille zum Zeitunglesen
Corrective focus:	<i>Maria nimmt die blaue Brille.</i> Maria nimm die [gelbe] <sub>foc</sub> Brille zum Zeitunglesen

↑  
Increase of prominence  
↓

Recall: lila – lila

→ increase of prosodic prominence



# Method – Data





# Method – Data annotation

File Edit Annotation Tier Type Search View Options Window Help

Grid Text Subtitles Lexicon Comments Recognizers Metadata Controls

Label

G01\_P\_07\_B · Pause · feedback · Pause · feedback · Pause · feedback · Pause · Geräusch · Pause · G01\_P\_03\_I · Pause · feedback · Pause · Feedback · Pause · G01\_P\_04\_F · Pause · feedback · Pause · feedback · Pause · G01\_P\_08\_B · Pause · feedback · Pause · feedback · Pause · Geräusch · Pause · feedback / andere Srtimmen · Pause · Geräusch · Pause · Geräusch · Pause · feedback · Pause · feedback · Pause · G01\_T\_01\_F · feedback · Pause · feedback · Pause · G01\_T\_02\_I · feedback · Pause · feedback · Pause · G01\_P\_05\_C · Pause · feedback · Pause · feedback · Pause · G01\_P\_06\_R · Pause · feedback · Pause · feedback · Pause · G01\_T\_29\_B · feedback · Pause · feedback · Pause · G01\_T\_03\_I · Pause · feedback · Pause · feedback · Pause · G01\_T\_04\_C · G01\_T\_05\_R · Pause · feedback · Pause · feedback · Pause · feedback · Pause · feedback · Pause · G01\_T\_30\_B · feedback · Pause · Geräusch · Pause · feedback · Pause · Geräusch · feedback · Pause · feedback · Pause · feedback · Pause · Geräusch · Pause · feedback · Pause · G01\_T\_07\_I · Pause · feedback · Pause

00:05:38.683 Selection: 00:00:00.000 – 00:00:00.000 0

Selection Mode Loop Mode

00:05:39.000 00:05:40.000 00:05:41.000 00:05:42.000 00:05:43.000 00:05:44.000 00:05:45.000 00:05:46.000

Satz [124] Um nach Paris reisen zu können brauchst du deinen grünen Koffer, den grünen Koffer. Genau, da hast d

Label [270] G01\_T\_01\_F feedback

FocCondition [44] Contrastive (fill) Focus

NP [54] NP

Wort [86] Adj N

Manual\_GUnit [31] G-unit

Man\_Articulator\_I [31] BHE-R

Man\_GPhase [125] preparation strok preparation stroke prep stroke recovery

Man\_Apex [53]

Head\_Movement\_Type [150] turn tilt tilt

Perc\_Gestural\_Prominen [66] 2 0

Perc\_Prosodic\_Prominen [66] 3 1

PitchAccents [65]

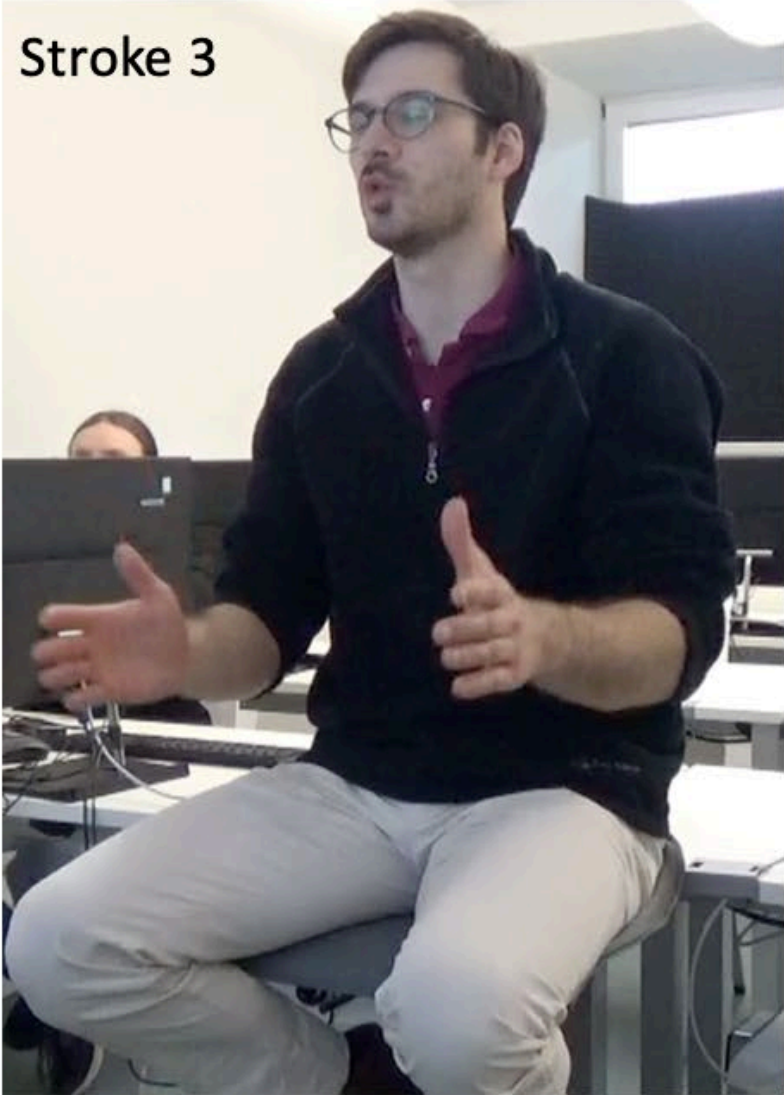
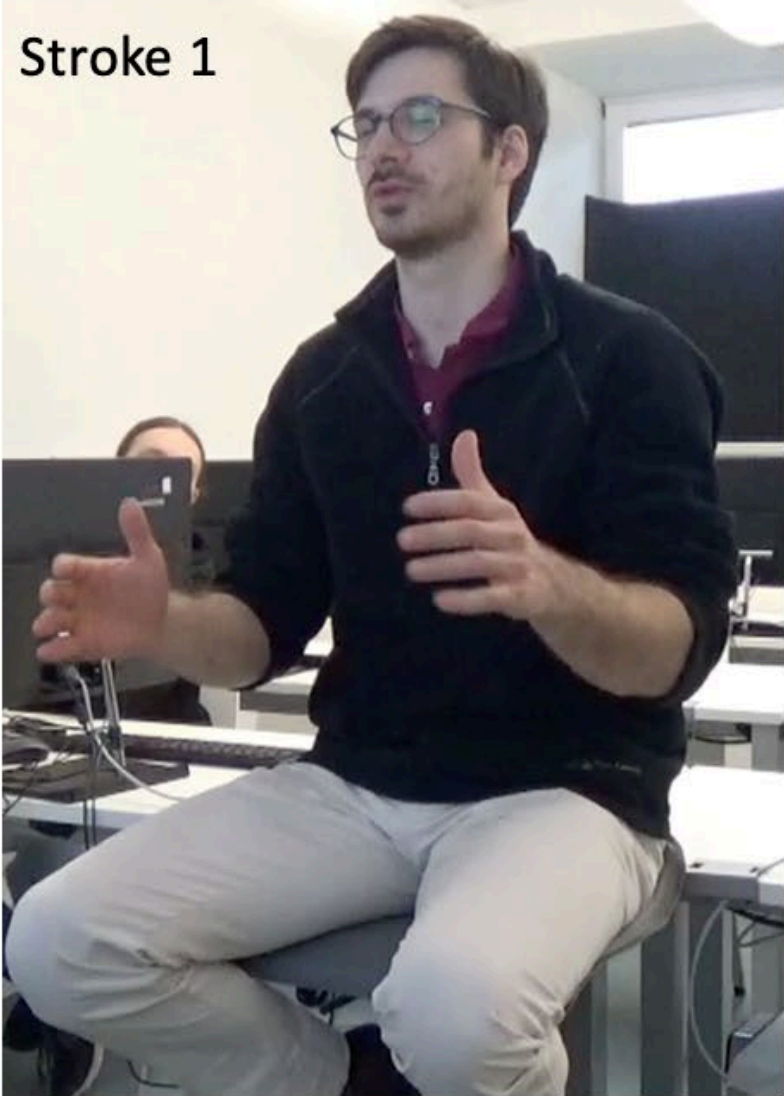
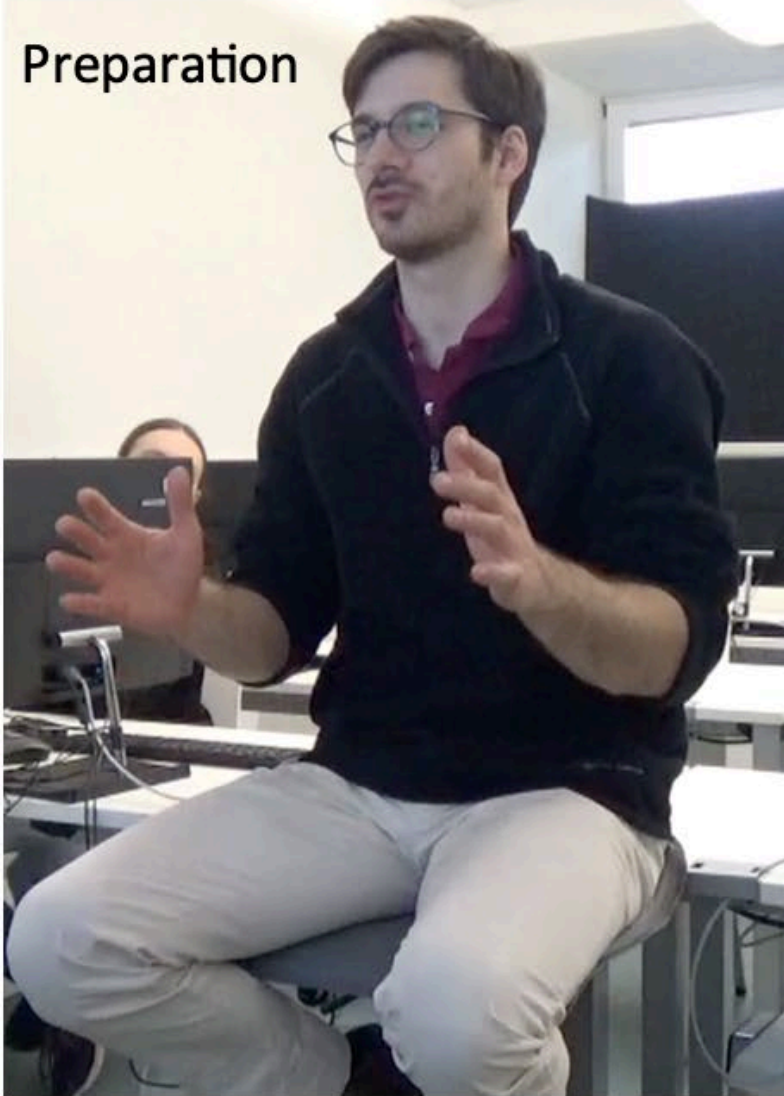
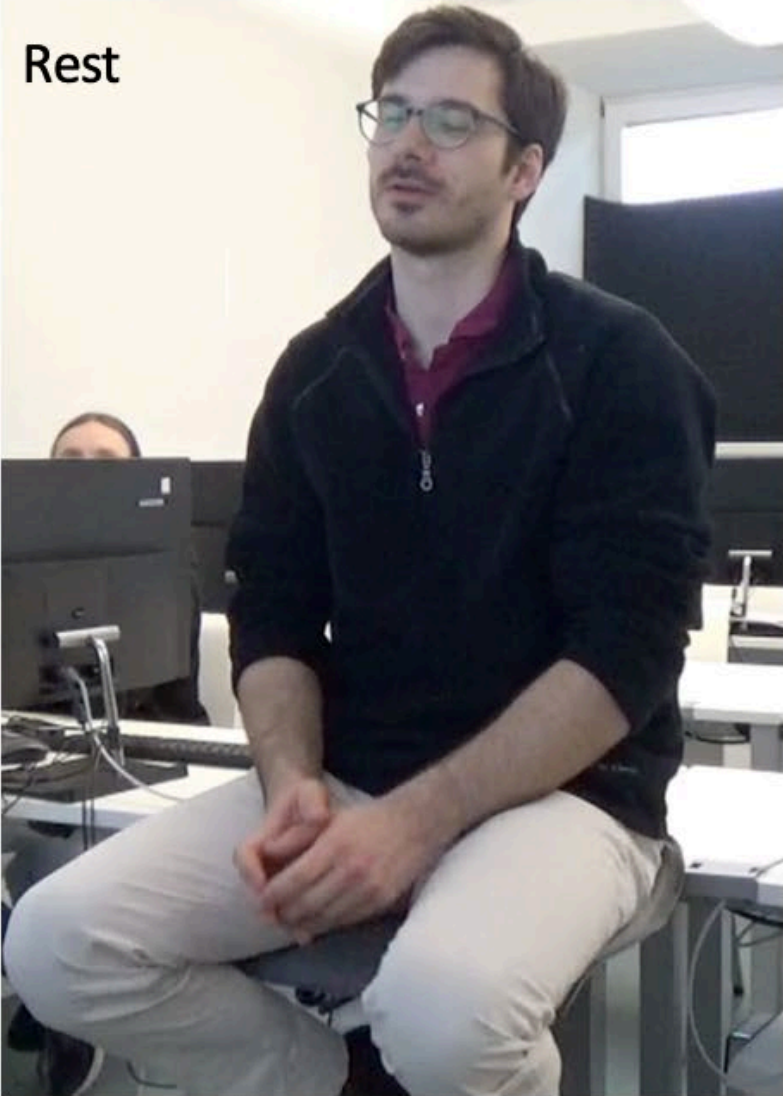
ProsodicPhrasingPlaceh [0]

Semantic\_id [52]

Non\_referential [46]

Iconicity [6] Iconi Iconic Iconic

# Inspection of a gesture



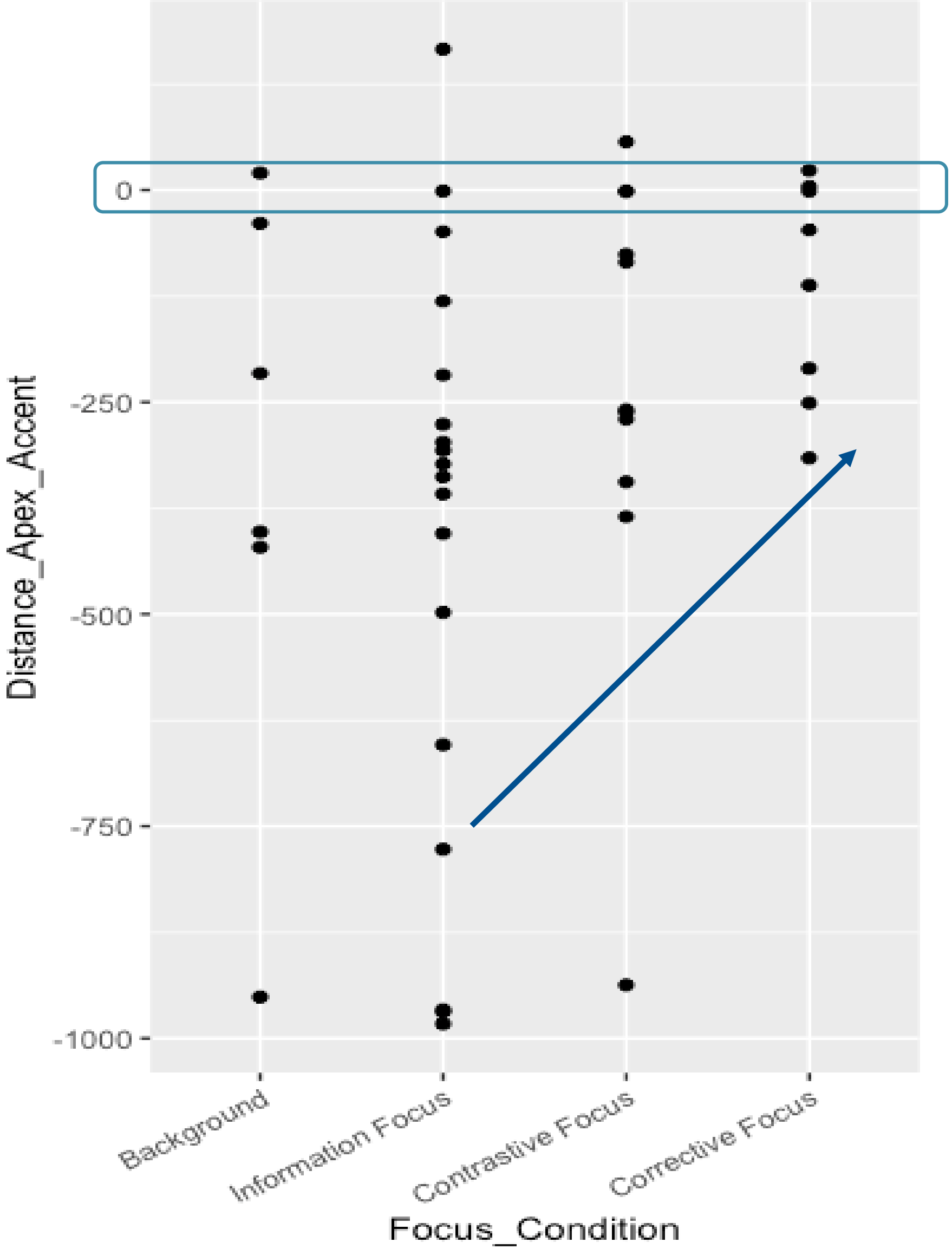
# Results – Overall inspection of gesture rate

Condition	Seconds (rounded)	Gestures	Gesture Rate (g/sec)
Background	170 sec	26	0,153
Information Focus	157 sec	35	0,223
Contrastive Focus	209 sec	62	0,3
Corrective Focus	156 sec	44	0,282
Total	703 sec	167	0,238

Increase of prominence

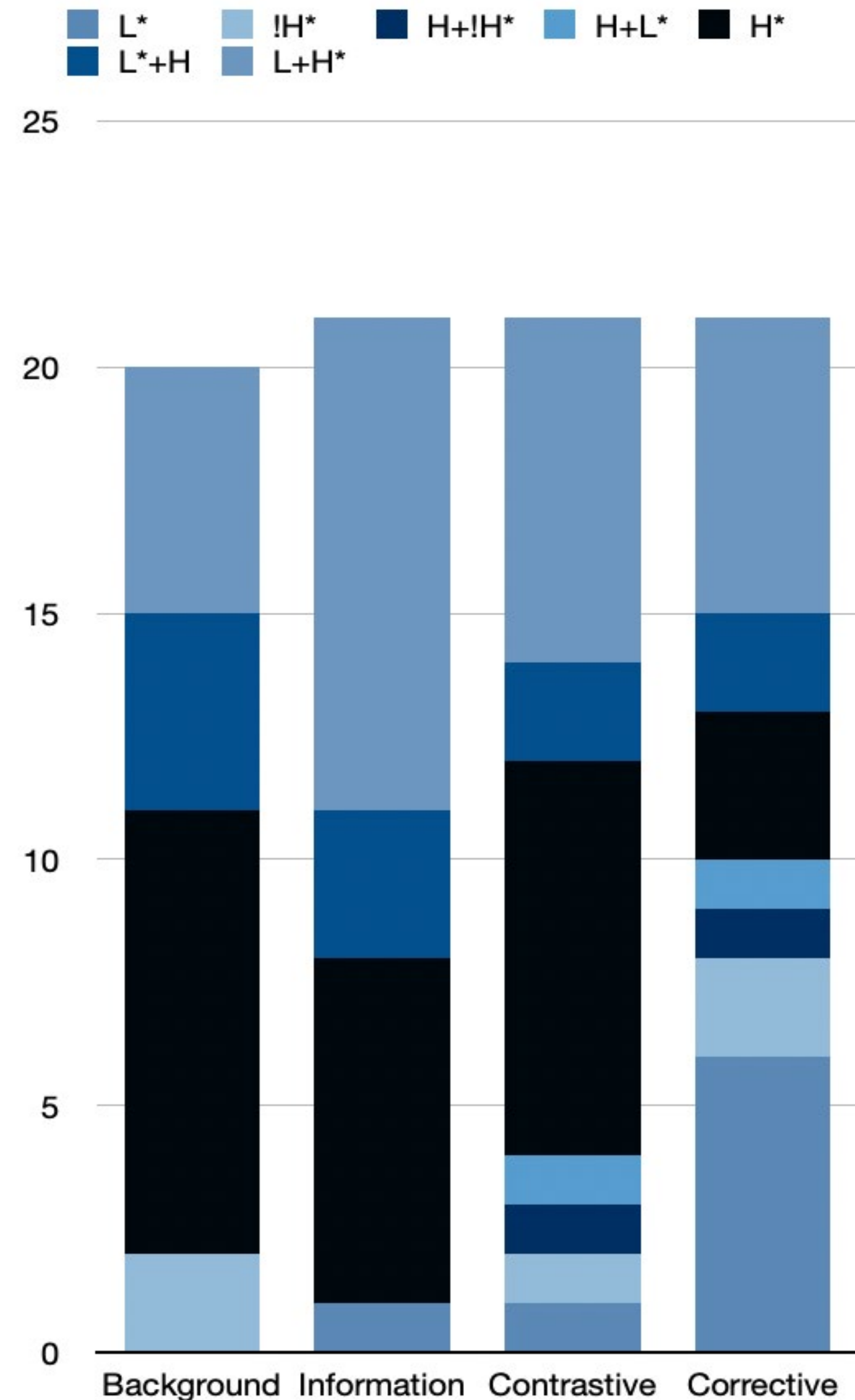
- Increase of gesture rate from less prominent to more prominent target sentences
- Measurement for the whole sentence!

# Results – Synchronisation of gesture and pitch accent



Increasing prominence ~ more precise gesture-to-speech coordination

# Results – Prosody, pitch accent types

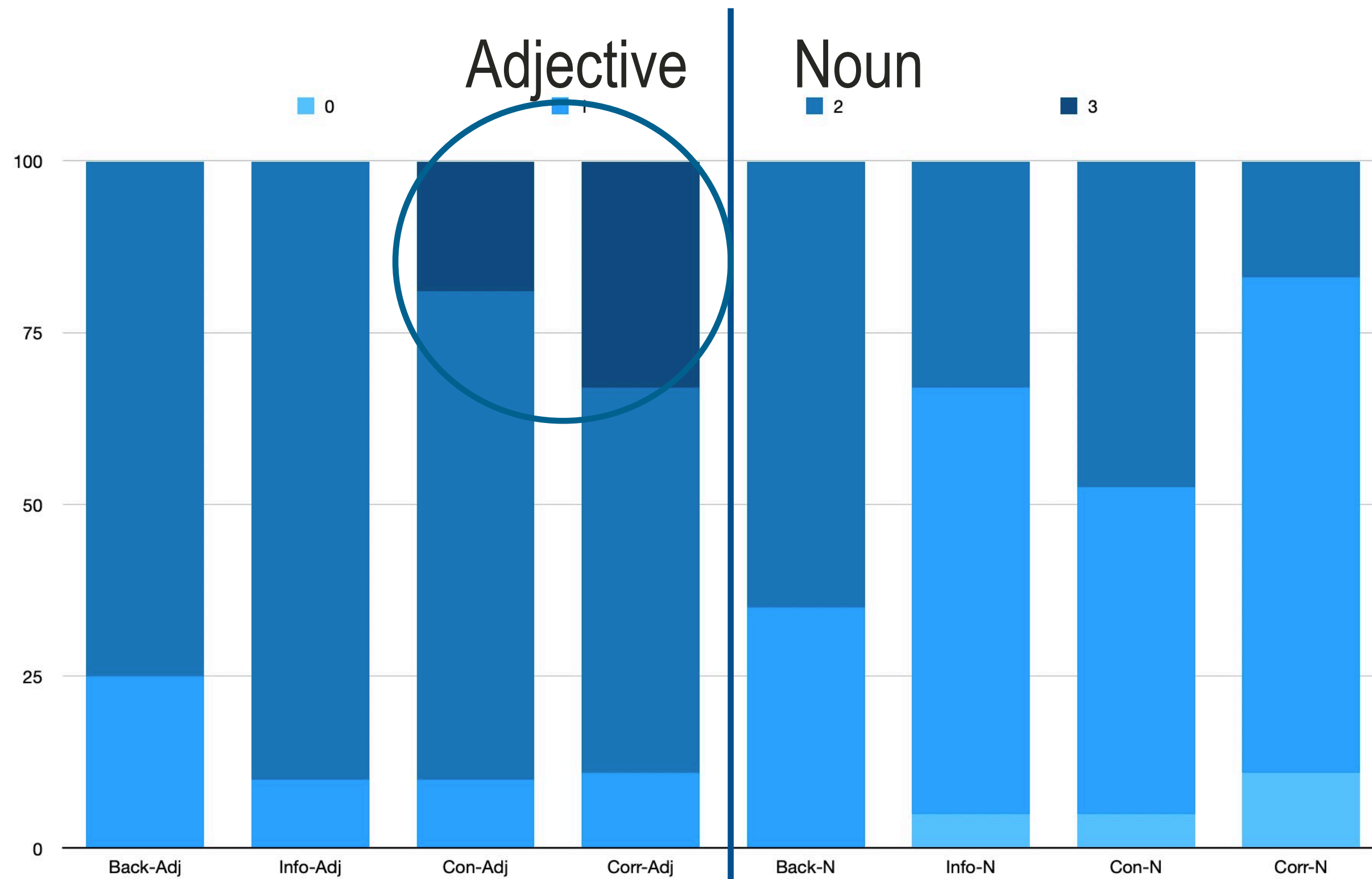


- Main pitch accent on the target phrase (adjective + noun)
- No clear pattern, except for all constituents are accented
- **In need of another measure for prosodic prominence!**

# Results – Prominence perception

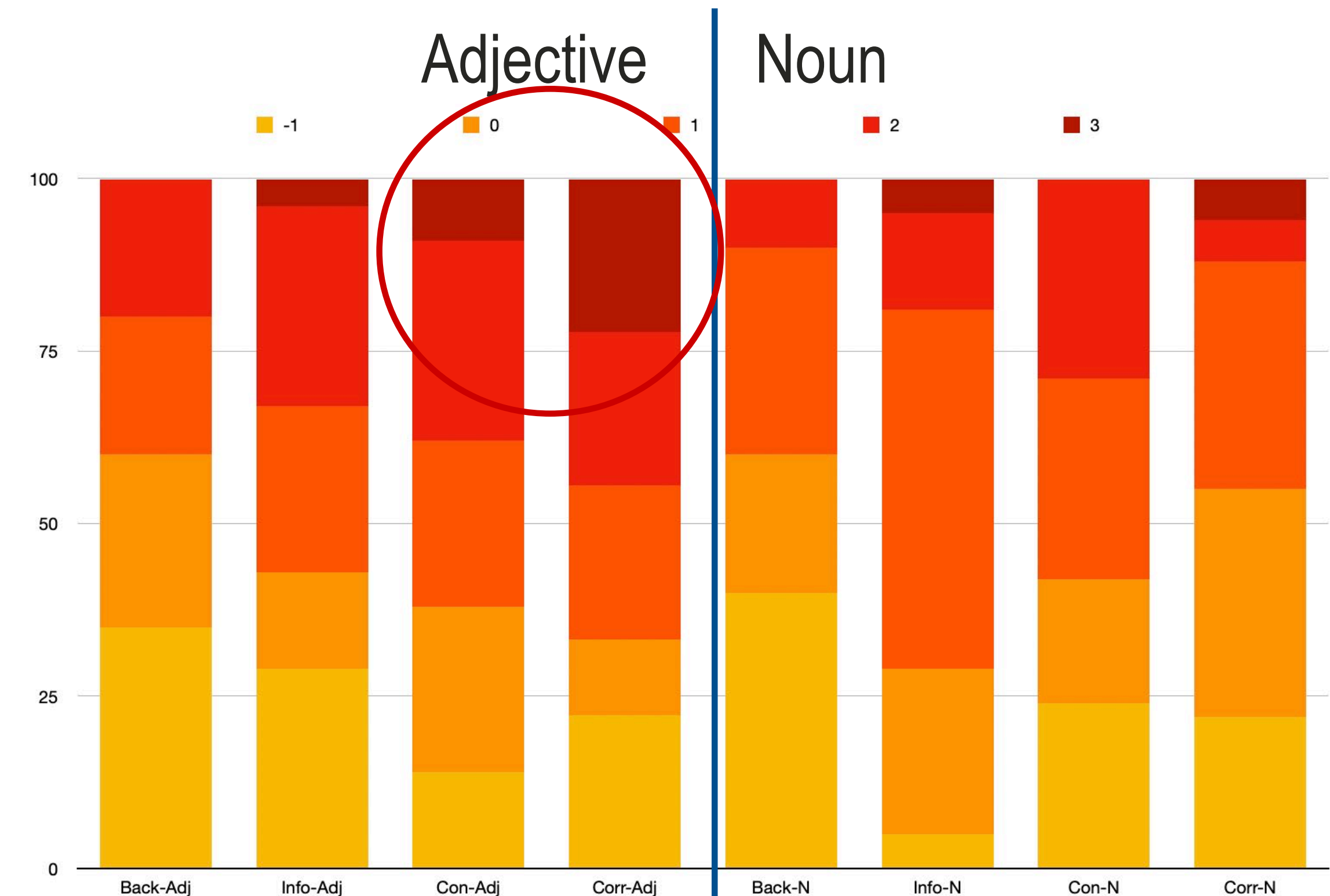
- Four prominence levels (according to DIMA, Kügler et al. 2022) – level 0 – 1 – 2 – 3
- Prosody perception in Praat (Boersma & Weenink 2023) without video; Gesture perception in ELAN without sound

## Prosodic prominence perception



Perceived prosodic (acoustic) prominence split by Focus Condition and Word

## Gestural prominence perception



Perceived visual prominence split by Focus Condition and Word

- An increase in prominence on adjective ~ increase in perceived prominence (both **acoustically** and **visually**)
- The noun is less prominent (both **acoustically** and **visually**)

# Discussion

Audio-visual prominence elicitation in the lab 😊

Global measure of gesture rate shows more gestures with higher prominence in the sentence.

Phonological pitch accent types are not informative

(despite assumed differences of prominence across different pitch accent types, cf. [Baumann & Röhr 2015](#)).

Perceived prominence of audio and video signal alone appear to show promising results.

- Increasing pragmatic prominence (context) leads to increased prosodic and gestural prominence.

# Summary & Conclusion

Corpus data of spontaneous speech (direction task) and experimentally elicited spontaneous speech show similar results.

Q1: Does prominence influence the occurrence of gestures in spontaneous German speech?

Yes! Corpus data: Increased occurrences of gestures in more prominent contexts.

Experimental data: Globally increased gesture rate (more gestures) in more prominent contexts.

➤ Prominence attracts co-speech gestures.

Q2: Is pitch accentuation temporally aligned with gesture apexes in German and does prominence influence this alignment?

Yes! Corpus data: More tight alignment of gestures with pitch accents in focus (prominent) than non-focus.

Experimental data: Increasing tighter alignment with increasing prominent contexts.

➤ Visual / gestural hyperarticulation ([Kügler & Gregori to appear](#))  
(similar to articulatory hyperarticulation; [Lindblom 1990](#))



# Summary & Conclusion

Corpus data of spontaneous speech (direction task) and experimentally elicited spontaneous speech show similar results.

Q3: Are different types of gestures (iconic and non-referential) affected by prominence in the same way and extent?

Yes! Corpus data: no difference in gesture-speech synchronization between iconic and non-referential gestures.

Experimental data: both iconic (Koffer 'suitcase') and non-referential (Löffel 'spoon') occur on prominent elements

➤ Confirmation of the multidimensional hypothesis of gestures ([McNeill 2006](#), [Rohrer et al. 2020](#))

Q4: Are different degrees of prominence encoded with different degrees multimodal prominence?

Yes! Experimental data: Increase of audio-visual prominence as a function of increased pragmatic prominence.

➤ Prominence attracts gestures, and increased prominence leads to more prominent gestures

➤ higher gestural prominence in terms of velocity, amplitude and “beat-like-ness”

# Take home message

**Prominence attracts prosodic and visual marking –  
the more pragmatic prominence the more prominent are acoustic and gestural cues**

**The gesture-prosody link is thus indirect – mediated by prominence**

# Thank You!

## Acknowledgements

Thanks to my colleagues Alina Gregori, Paula Sánchez-Ramón and Pilar Prieto.

Thanks to Andy Lücking for providing the SaGA corpus.

Thanks to the DFG and Goethe University Frankfurt for funding this research.

This work is part of the project **MultIS** in the DFG priority programme SPP 2392 Visual Communication (**ViCom**).

contact: [kuegler@lingua.uni-frankfurt.de](mailto:kuegler@lingua.uni-frankfurt.de)

- Ambrazaitis, Gilbert; House, David (2022): Phonological Categories: Identification, representation, implementation. In: *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 24 (1). DOI: 10.16995/labphon.6430.
- Baumann, Stefan; Röhr, Christine T. (2015): The perceptual prominence of pitch accent types in German. In: *The Scottish Consortium for ICPHS 2015 (Hg.): Proceedings of the 18th International Congress of Phonetic Sciences. International Congress of Phonetic Sciences. Glasgow, 2015. Glasgow, UK: University of Glasgow, S. 384. Online verfügbar unter <http://www.icphs2015.info/pdfs/Papers/ICPHS0384.pdf>.*
- Boersma, Paul; Weenink, David (2023): Praat: doing phonetics by computer [Computer program]. Version 6.3.10 (<http://www.praat.org/>). Version Version 6.3.10. Online verfügbar unter <http://www.praat.org/>.
- Esteve-Gibert, Núria; Prieto, Pilar (2013): Prosodic Structure Shapes the Temporal Realization of Intonation and Manual Gesture Movements. In: *J Speech Lang Hear Res* 56 (3), S. 850–864. DOI: 10.1044/1092-4388(2012/12-0049).
- Féry, Caroline (2017): *Intonation and Prosodic Structure*. Cambridge: Cambridge University Press (Key Topics in Phonology). Online verfügbar unter <http://ebookcentral.proquest.com/lib/gbv/detail.action?docID=4786311>.
- Féry, Caroline; Kügler, Frank (2008): Pitch accent scaling on given, new and focused constituents in German. In: *Journal of Phonetics* 36 (4), S. 680–703.
- Götze, Michael; Weskott, Thomas; Endriss, Cornelia; Fiedler, Ines; Hinterwimmer, Stefan; Petrova, Svetlana et al. (2007): Information structure. In: Stefanie Dipper, Michael Götze und Stavros Skopeteas (Hg.): *Information Structure in Cross-Linguistic Corpora. Annotation Guidelines for Phonology, Morphology, Syntax, Semantics, and Information Structure*. Potsdam: Universitätsverlag Potsdam (Interdisciplinary Studies on Information Structure (ISIS), 7), S. 147–187.
- Gregori, Alina (2022): *Co-speech Gestures, Information Structure and Prosody. A Corpus Study on Prominence Peak Alignment*. MA thesis. Goethe Universität Frankfurt, Frankfurt am Main.
- Gregori, Alina; Amici, Federica; Brilmayer, Ingmar; Ćwiek, Aleksandra; Fritzsche, Lennart; Fuchs, Susanne et al. (2023): A Roadmap for Technological Innovation in Multimodal Communication Research. In: *Proceedings 25th International Conference on Human-Computer Interaction. Conference on Human-Computer Interaction. Copenhagen, Denmark, July 2023*. Berlin: Springer, S. 1–38.
- Grice, Martine; Baumann, Stefan; Benz Müller, Ralf (2005): German Intonation in Autosegmental-Metrical Phonology. In: Sun-Ah Jun (Hg.): *Prosodic Typology. The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press, S. 55–83.
- Grice, Martine; Kügler, Frank (2021): Prosodic Prominence - A Cross-Linguistic Perspective. In: *Language and Speech* 64 (2), S. 253–260. DOI: 10.1177/00238309211015768.
- Gussenhoven, Carlos (2004): *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press.
- Gussenhoven, Carlos; Chen, Aoju (2020a): Introduction. In: Carlos Gussenhoven und Aoju Chen (Hg.): *The Oxford Handbook of Language Prosody*. Oxford: Oxford University Press, Ivi-12.
- Gussenhoven, Carlos; Chen, Aoju (Hg.) (2020b): *The Oxford Handbook of Language Prosody*. Oxford: Oxford University Press.
- Hanssen, Judith; Peters, Jörg; Gussenhoven, Carlos (2008): Prosodic Effects of Focus in Dutch Declaratives. In: P. A. Barbosa, S. Madureira und C. Reis (Hg.): *Proceedings of Speech Prosody 2008 Conference*. Campinas, Brazil: Editora RG/CNPq, S. 609–612.
- Im, Suyeon; Baumann, Stefan (2020): Probabilistic relation between co-speech gestures, pitch accents and information status. In: *Proceedings of the Linguistic Society of America* 5 (1), S. 685–697.
- Kendon, Adam (1980): Gesticulation and speech: Two aspects of the process of utterance. In: Mary R. Key (Hg.): *The Relationship of Verbal and Nonverbal Communication*. Berlin: Mouton, S. 207–227.
- Kendon, Adam (2004): *Gesture. Visible action as utterance*. Cambridge: Cambridge University Press. Online verfügbar unter <http://www.loc.gov/catdir/description/cam041/2004049686.html>.
- Krifka, Manfred (2008): Basic notions of information structure. In: *Acta Linguistica Hungarica* 55 (3), S. 243–276.

- Kügler, Frank; Baumann, Stefan; Röhr, Christine T. (2022): Deutsche Intonation, Modellierung und Annotation (DIMA). Richtlinien zur prosodischen Annotation des Deutschen. In: Cordula Schwarze und Sven Grawunder (Hg.): Transkription und Annotation gesprochener Sprache und multimodaler Interaktion. Konzepte, Probleme, Lösungen. Tübingen: Narr Francke Attempto Verlag, S. 23–54.
- Kügler, Frank; Calhoun, Sasha (2020): Prosodic Encoding of Information Structure. A typological perspective. In: Carlos Gussenhoven und Aoju Chen (Hg.): The Oxford Handbook of Language Prosody. Oxford: Oxford University Press, S. 453–467.
- Kügler, Frank; Gollrad, Anja (2015): Production and Perception of Contrast. The case of the rise-fall contour in German. In: *Frontiers in psychology* 6 (1254), S. 1–18. DOI: 10.3389/fpsyg.2015.01254.
- Kügler, Frank; Gregori, Alina (2023): Iconic Gestures in Focus - Synchronization of Prosody and Gestures in Prominence. In: *ICPhS* (Hg.): Proceedings of ICPhS 20. Prague. Prague, S. 1–5.
- Ladd, D. Robert (1996): *Intonational Phonology*. Cambridge: Cambridge University Press.
- Ladd, D. Robert (2008): *Intonational Phonology*. 2. Aufl. Cambridge: Cambridge University Press.
- Lindblom, Björn (1990): Explaining phonetic variation. A sketch of the H&H theory. In: William J. Hardcastle und A. Marchal (Hg.): *Speech Production and Speech Modelling*. Dordrecht: Kluwer, S. 403–439.
- Loehr, Daniel P. (2012): Temporal, structural, and pragmatic synchrony between intonation and gesture. In: *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 3 (1), S. 71–89. DOI: 10.1515/lp-2012-0006.
- Lücking, Andy; Bergmann, Kirsten; Hahn, Florian; Kopp, Stefan; Rieser, Hannes (2010): The Bielefeld Speech and Gesture Alignment Corpus (SaGA). In: M. Kipp, J.-P. Martin, P. Paggio und D. Heylen (Hg.): *LREC 2010 Workshop. Multimodal Corpora—Advances in Capturing, Coding and Analyzing Multimodality, 2010*. ELRA, S. 92–98.
- McNeill, David (1992): *Hand and mind. What gestures reveal about thought*. Chicago: University of Chicago Press.
- McNeill, David (2006): Gesture. a psycholinguistic approach. In: Edward Keith Brown (Hg.): *Encyclopedia of language and linguistics*. 2nd ed. Amsterdam: Elsevier, S. 58–66.
- Nespor, Marina; Vogel, Irene (1986): *Prosodic phonology*. Dordrecht: Foris Publ. (Studies in Generative Grammar, 28).
- Peters, Jörg (2014): *Intonation*. 1. Aufl. Heidelberg: Universitätsverlag Winter (Kurze Einführungen in die germanistische Linguistik - KEGLI, 16).
- Pierrehumbert, Janet B. (1980): *The phonology and phonetics of English intonation*. PhD Thesis. MIT, Cambridge, Massachusetts.
- Rohrer, Patrick L. (2022): *A temporal and pragmatic analysis of gesture-speech association. A corpus-based approach using the novel MultiModal MultiDimensional (M3D) labeling system*. PhD thesis. Universitat Pompeu Fabra, Barcelona.
- Rohrer, Patrick Louis; Vilà-Giménez, Ingrid; Florit-Pons, Júlia; Gurrado, G.; Esteve-Gibert, Núria; Ren, Ada et al. (2020): *The MultiModal MultiDimensional (M3D) labeling system for the annotation of audiovisual corpora: Gesture Labeling Manual*. UPF Barcelona.
- Sánchez-Ramón, Paula G.; Gregori, Alina; Prieto, Pilar; Kügler, Frank (2023): The impact of focus types on the prosody-gesture link in Catalan and German: a focus elicitation production study. 1st International Multimodal Communication Symposium. MMSYM. Barcelona, April 2023.
- Selkirk, Elisabeth O. (1984): *Phonology and syntax. The relation between sound and structure*. Cambridge, MA: MIT Press.
- Shattuck-Hufnagel, Stefanie; Yasinnik, Yelena; Veilleux, Nanette; Renwick, Margaret (2007): A method for studying the time-alignment of gestures and prosody in American English. 'Hits' and pitch accents in academic-lecture-style speech. In: Anna Esposito, Maja Bratanic, Eric Keller und Maria Marinaro (Hg.): *Fundamentals of verbal and nonverbal communication and the biometric issue*. Amsterdam: IOS Press, S. 34–44.