

VII Falas no LabFon  
28 de Março

**Perceiving incredulity:  
the role of intonation and facial  
expression**

Verónica Crespo-Sendra

Laboratório de Fonética  
Faculdade de Letras, Universidade de Lisboa

# Introduction

- **Recent studies have demonstrated that communicative functions signaled by auditory prosody tend to be supported by visual information as well as in the form of specific facial expressions.**  
(Krahmer & Swerts 2005, Dijkstra, Krahmer & Swerts 2006, Borràs-Comes & Prieto 2010, among others)
- **Some studies have shown that visual information partially duplicates the information of auditory cues and facilitates the decoding process.**  
(Srinivasan & Massaro 2003, House 2002, Dohen & Løevenbruck 2009)
- **Yet, little is known about the relative contributions of gestures and intonation information in the decoding of linguistic meaning.**

# Introduction

- To gain more insight into possible relationships between auditory and visual cues, this study looks into the audiovisual prosodic marking of questions.
- Little is known about the interaction between visual and acoustic cues in the conveyance of pragmatic information in questions.

# Introduction

## Previous research on questions: two perspectives

➤ **Some studies have found that influence of the visual cues on the auditory cues was marginal.**

(Srinivasan & Massaro (2003) for questions vs. statements in English, House (2002), for questions vs. statements in Swedish)

➤ **Other studies have revealed a consistent and strong effect of visual cues in the listener's decisions, but also a consistent effect of the auditory stimulation.**

(Borràs-Comes & Prieto (2010), Borràs-Comes, Pugliesi & Prieto (2011), echo questions vs. contrastive focus in Catalan)

# Introduction

## The Fuzzy Logical Model of Perception (FLMP)

**These studies were consistent with the FLMP, which predicts that:**

- both auditory and visual modalities will influence the perception of prosody,
- the influence of one modality will be greater when the information provided by the other modality is ambiguous.

(Srinivasan & Massaro 2003)

# Justification

- **Contradictory results in the literature.**
- **Lack of cross-linguistic studies investigating the differences in the weight of gestures and acoustic cues.**

Clearly, there is a need for a **better understanding** of the semantic weight carried by the gestural cues, as well as the **interaction** between the contributions of the **intonation** and **facial gestures** of the sentences in conveying different pragmatic meanings.

# Justification

**Goal:** to investigate the interaction between facial gestures and intonation on linguistic questions interpretation.

**Subject of study:** two languages, Catalan and Dutch.



# Justification

**Type of questions:** Information-seeking questions vs. Incredulity questions (those that convey *mirativity*).

From a typological perspective the linguistic marking of new or unexpected information of the speakers (which has been called *mirativity* in the pragmatics literature) can be conveyed through the intonation system in some languages, and in some others through the morphosyntactic system (DeLancey, 2001).



**Why:** both languages are expected to mark the difference between the two question types through different intonational contours and both languages seem to use similar visual cues for this contrast.



# Production Experiment

In order to confirm the intonation of the two question types and which gestural patterns be used in both languages, 8 native speakers of Catalan (ages 20-40) and 8 native speakers of Dutch were videotaped pronouncing both possible interpretations of the utterance.

Participants were asked to answer in an expressive way two context with one word (a short utterance might make the identification a more automatic perceptual task, and less of a cognitive decision-making process, Srinivasan & Massaro 2003).

# Production Experiment

## Info-seeking

–You are talking with your partner about the party that you are organizing for tonight. He is talking about the dessert, and after saying that you already have apples, bananas and pears, he asks you what else you could buy. You make a tentative proposal in the form of a question about tangerines.

(Recorded response: *Tangerines?*)



## Incredulity

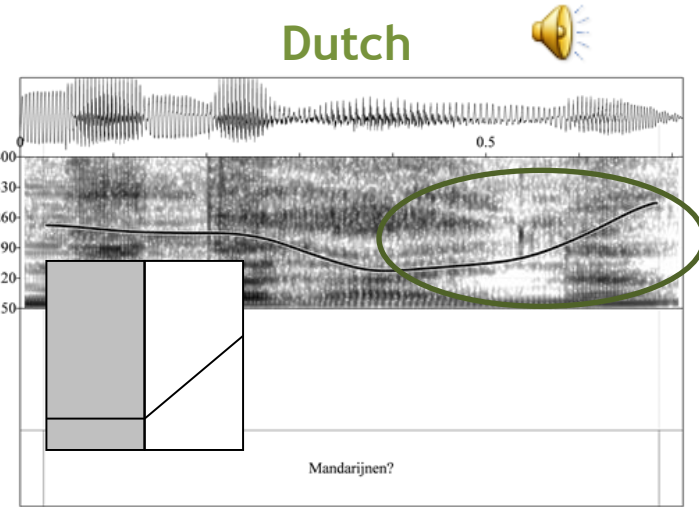
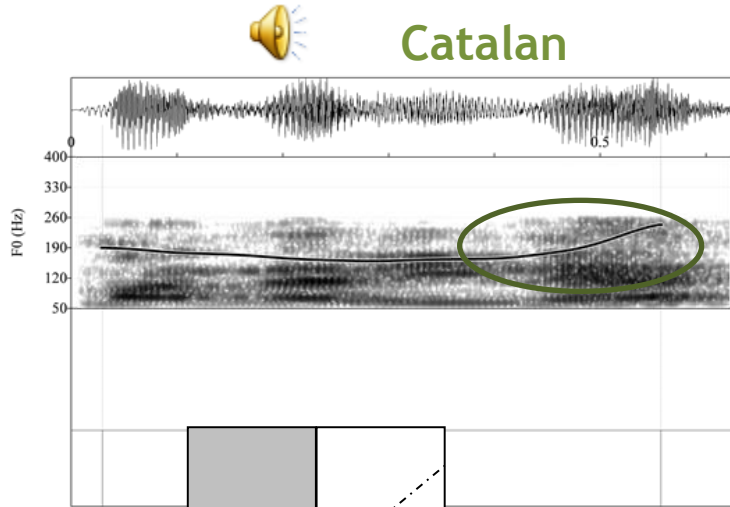
–You enter your home and see that your mother is eating tangerines. You know that your mother doesn't like tangerines. Unable to believe your eyes, you ask her about the tangerines.

(Recorded response: *Tangerines!?*)

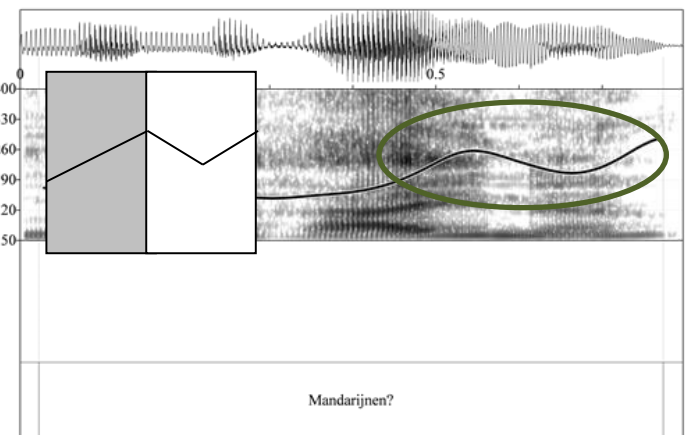
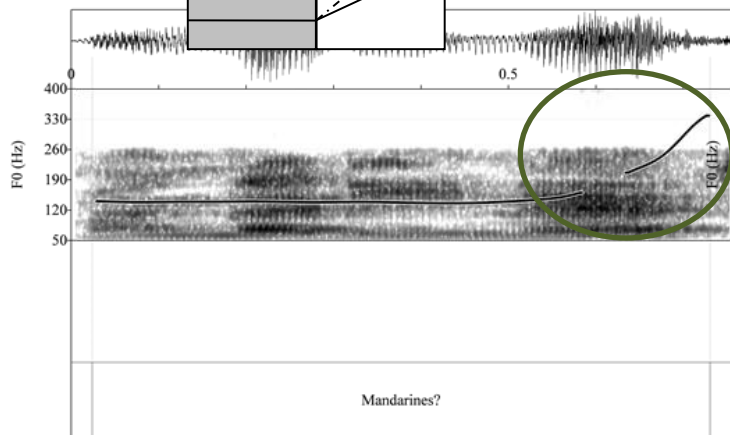


# Production Experiment / Intonation

Information-seeking question



Counter-expectational/  
Incredulity question



Pitch range differences vs.

Pitch accent+boundary tone

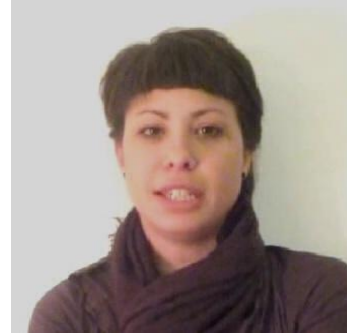


# Production Experiment / Gestures

## Information-seeking question

- ✓ upward eyebrow movement
- ✓ head down and head up

Catalan



Dutch



## Counter-expectational/ Incredulity question

- ✓ downward eyebrow movement
- ✓ squinting of the eyes
- ✓ head down and head back



# Goal / Hypothesis

## GOAL

This contrast constitutes a unique opportunity to test whether listeners from these two languages can use different processing strategies in the course of semantic interpretation.

## HYPOTHESIS

According to the **FLMP** theory, we hypothesize that Catalan listeners will rely more heavily on visual information since acoustic information is weaker than gestural patterns; while Dutch speakers will rely more on the strong gestural patterns of the intonational contrast.

# Perception Experiment

## Modalities

- Audio Only (AO)
- Video Only (VO)
- Audiovisual (AV)

## Materials

- 2 audio files (neutral, incredulity) x 5 speakers
  - 2 video files (neutral, incredulity) x 5 speakers
  - 2 audiovisual files (neutral, incredulity) x 2 situations (congruous, incongruous) x 5 speakers
- All of them were **original stimuli from 5 native speakers of each language.**

## Instructions

- Participants had to answer whether they interpret the audio, visual and AV stimuli as “neutral” or “incredulous” (N or I)
- 20 native speakers of each language participated
- Experiment set up by means of E-Prime 2



# Perception Experiment

## Audio Only

2 audio files (informative, incredulity) x 5 speakers x 3 blocks x 20 participants:  
**600** responses per language



## Video Only

2 video files (informative, incredulity) x 5 speakers x 3 blocks x 20 participants:  
**600** responses per language



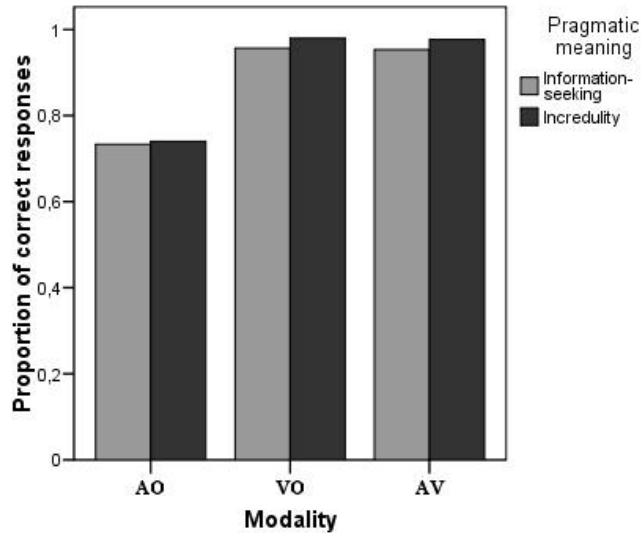
## Audiovisual

2 audiovisual files (informative, incredulity) x 2 situations (congruous, incongruous) x 5 speakers x 3 blocks x 20 participants: **1200** responses per language



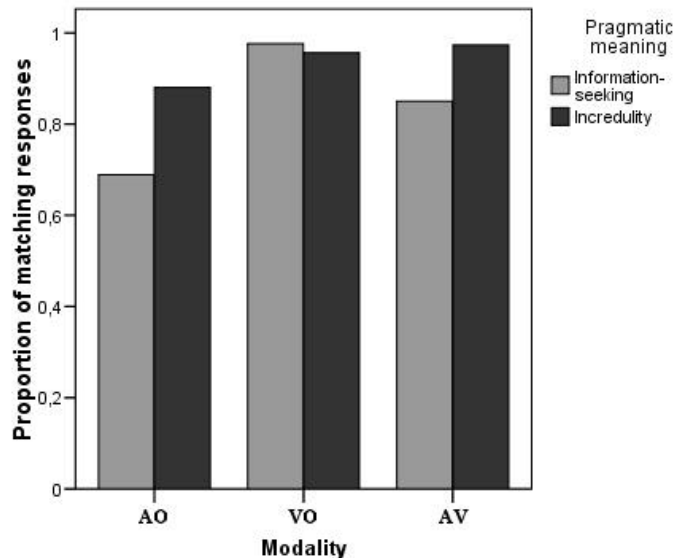
All the stimuli were extracted from 5 original videos of each language by Adobe Premiere (AdobeSystems Incorporated 2002)

# Results (congruent stimuli)



## Catalan

**AO** stimuli were not very strong. By contrast, results of the **VO** and **AV** stimuli demonstrate that visual cues were basic in determining speakers' decision.



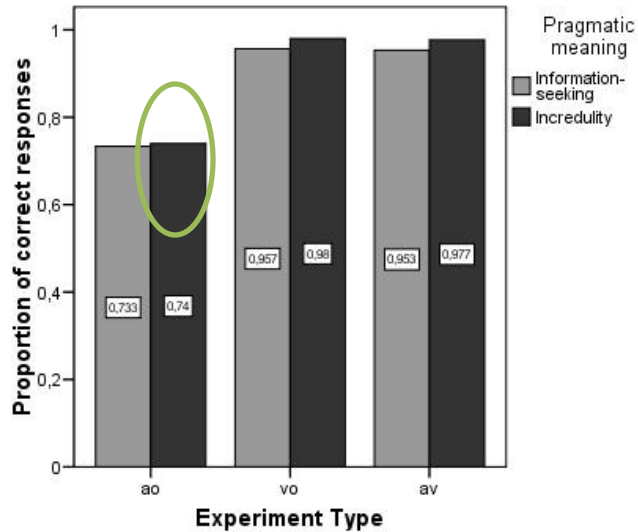
## Dutch

**AO** incredulity stimuli is stronger  
**VO** task are quite similar in both pragmatic meanings,  
**AV** speakers are able to correctly identify the incredulity cues but in the information-seeking questions there is a lower proportion of matching responses.



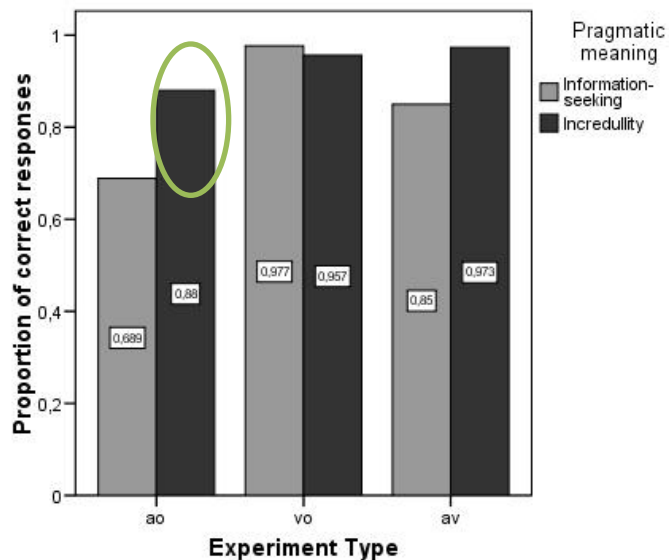
# Results (congruent stimuli)

## Catalan



Dutch audio stimuli incredulity are more efficient than in Catalan, because of the different configuration, but *what happens when we have incongruent stimuli?*

## Dutch



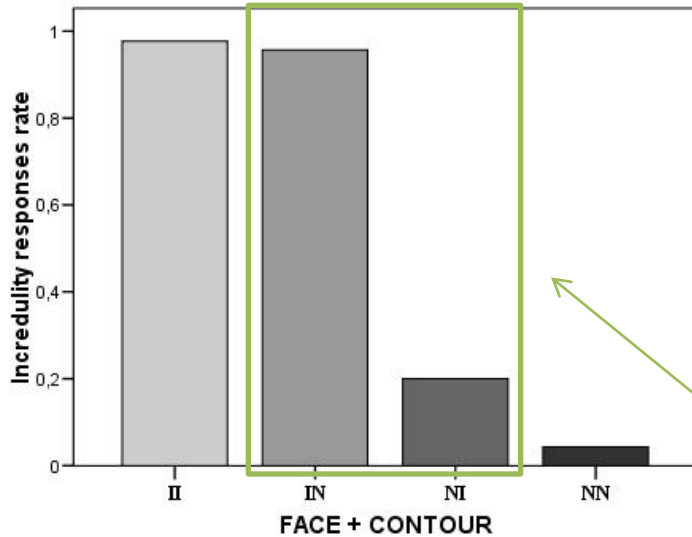
# Results AV (congruent , incongruent stimuli)

## Catalan

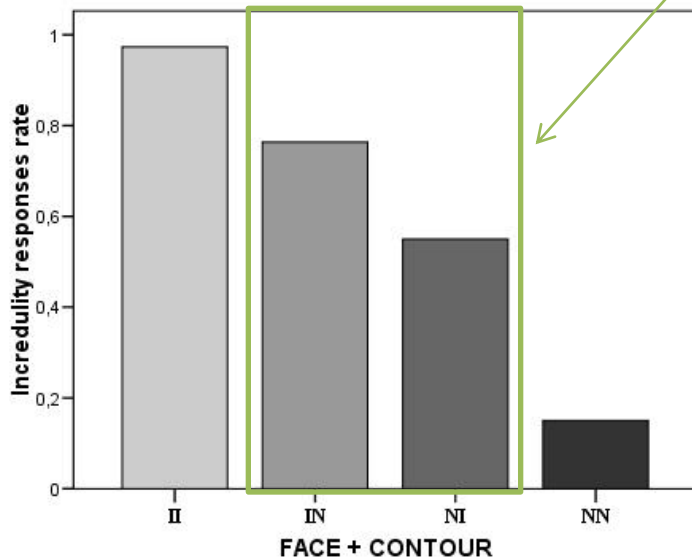
This means that given an incongruent situation, listeners **rely** heavily on the **facial expression** in distinguishing both question types.

## Dutch

The perceptual ratings of incongruent stimuli appear to be strongly affected by the **gestural and prosodic cues to incredulity**, Dutch speakers seem to rely strongly on facial cues, but there is evidence that these cues are not enough to obtain a 100% correct responses, and they have a clear influence of intonation contours.



## Incongruent stimuli



# Results Summary

- ✓ Results from identification task in three sensory modalities (AO, VO, AV) demonstrate that there are **differences** between Catalan and Dutch listeners **in the perceptual processing of incredulity questions**.
  - In general, visual cues (that is, in VO or AV modalities) play a stronger role than auditory cues (AO) in both languages and induce higher degrees of correct identification.
  - AO: the specific incredulity contour in Dutch induces an increase of 30% in correct identification responses with respect to Catalan, a fact that could be explained by the **specific and marked contour type of the incredulity questions** in Dutch.
- ✓ Results of the **AV** condition with **congruent** and **incongruent** materials demonstrate that the congruent stimuli achieve higher identification scores.
  - Yet for the identification of **incongruent stimuli**, while **Dutch** listeners' perceptual ratings are strongly affected by both **gestural** and **prosodic** cues to **incredulity**, **Catalan** listeners' rely more strongly on **facial expression**.

# Conclusions

- The results agree with the **FLMP** model of speech perception; i.e., if one cue is ambiguous or weaker in one modality enhances the role of the other modality.
  - Thus, in **Catalan**, decisions are strongly affected by **facial expression**, because the difference in intonation between the two question types is less salient (i.e., they have the same nuclear configuration).
  - By contrast, in **Dutch**, the strongest cues are both the **incredulity facial expression** and the **incredulity pitch contour**; in other words, listeners identify a neutral facial expression with an incredulity intonation contour as incredulous, and they also perceive an incredulous meaning when the stimuli are an incredulous facial expression combined with a neutral intonation contour.

# Conclusions

- In general, the results of this set of experiments reveal the importance of the visual cues in the identification of tune meaning, thus confirming the results of recent studies on **audiovisual prosody** (Swerts & Kraemer 2005; Borràs & Prieto 2010, among others).
- Moreover, the results of the present study show that two different languages that express **mirativity** in a prosodic fashion assign different weights (in both perception and production) to visual and intonational cues.
- Finally, we believe that the results presented here support the idea that gestures and intonation must be investigated together when the perception of a given meaning is analyzed.

Muito obrigada!

*Moltes gràcies!*

# References

- Borràs-Comes, Joan; Prieto, Pilar (2010). "Seeing tunes". The role of visual gestures in tune interpretation. *12th Conference on Laboratory Phonology (LabPhon 12)*. University of New Mexico.**
- DeLancey, S. (2001) The mirative and evidentiality. *Journal of Pragmatics*, 33, pp. 369-382.**
- Dijkstra, C., Krahmer, E., & Swerts, M. (2006). Manipulating Uncertainty: The contribution of different audiovisual prosodic cues to the perception of confidence. *Proceedings of the Third International Conference on Speech Prosody*, Dresden.**
- Dohen, M., & Løevenbruck, H. (2009). Interaction of audition and vision for the perception of prosodic contrastive focus. *Language and Speech*, 52(2/3), 177-206.**
- House, D. (2002). Intonation and visual cues in the perception of interrogative**
- Krahmer, E., & Swerts, M. (2005). How children and adults produce and perceive uncertainty in audiovisual speech. *Language and Speech*, 48(1), 29-54.**
- Srinivasan, R. J., & Massaro, D. W. (2003). Perceiving from the face and voice: Distinguishing statements from echoic questions in English. *Language and Speech*, 46(1), 1-22.**