



## **Discours**

Revue de linguistique, psycholinguistique et informatique. A journal of linguistics, psycholinguistics and computational linguistics

**13 | 2013**  
**Varia**

---

# Syntactic and Prosodic Forms of First Names in Institutional Interaction Involving Multiple Participants

Dimitri Voilmy and Mari Wiklund

---



### **Electronic version**

URL: <http://journals.openedition.org/discours/8869>

DOI: 10.4000/discours.8869

ISSN: 1963-1723

### **Publisher:**

Laboratoire LATTICE, Presses universitaires de Caen

### **Electronic reference**

Dimitri Voilmy and Mari Wiklund, « Syntactic and Prosodic Forms of First Names in Institutional Interaction Involving Multiple Participants », *Discours* [Online], 13 | 2013, Online since 23 December 2013, connection on 20 April 2019. URL : <http://journals.openedition.org/discours/8869> ; DOI : 10.4000/discours.8869

---



*Discours* est mis à disposition selon les termes de la licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International.





Revue de linguistique, psycholinguistique et informatique

<http://discours.revues.org/>

## Syntactic and Prosodic Forms of First Names in Institutional Interaction Involving Multiple Participants

---

Dimitri Voilmy

Institut Charles Delaunay CNRS UMR 6279  
Université de technologie de Troyes

Mari Wiklund

Département des langues modernes  
Université de Helsinki, Finlande

.....  
Dimitri Voilmy et Mari Wiklund, « Syntactic and Prosodic Forms of First Names in Institutional Interaction Involving Multiple Participants », *Discours* [En ligne], 13 | 2013, mis en ligne le 23 décembre 2013.

.....  
URL : <http://discours.revues.org/8869>

.....  
Titre du numéro : *Varia*

Coordination : Stéphane Ferrari et Christian Raymond

**revues.org**  
CENTRE POUR L'ÉDITION ÉLECTRONIQUE OUVERTE  
CENTRE FOR OPEN ELECTRONIC PUBLISHING

 discours

 Presses  
universitaires  
de Caen



# Syntactic and Prosodic Forms of First Names in Institutional Interaction Involving Multiple Participants

---

Dimitri Voilmy

Institut Charles Delaunay CNRS UMR 6279  
Université de technologie de Troyes

Mari Wiklund

Département des langues modernes  
Université de Helsinki, Finlande

.....

This paper examines one aspect of turn-taking organization in institutional interactions: the use of first names and their prosodic marking for next-speaker selection. Institutional interaction is characterized by asymmetrical rights to talk and pre-allocation of action. This involves the restriction of one party to asking questions and the other to responding to them. The analysis focuses on two of these multiparty formal situations: co-present classroom participants and live interactive television broadcast with remote participants. In each context, turn allocation is determined by one party: the teacher or TV host. After asking a question as a sequence-initiating action, the teacher or host designates the next speaker by name. The use of first names is situatedly examined in terms of turn-taking organization and prosodic characteristics. The study examines how the prosodic marking is context-sensitive: do the participants have visual access to each other's actions and how is a name used to attract attention? This paper analyses the formation and maintaining of a mutual orientation towards a single conversational action: selecting and giving the floor to a co-participant of the conversation in an institutional framework. These detailed descriptions of the sequential order are based on ethnomethodologically-informed conversation analysis. The objective is to compare four "single cases", preserving the specificities and "whatness" of each excerpt.

**Keywords:** turn-taking, prosodic production, multiparty settings, conversation analysis, interactional linguistics, classroom interaction, media interaction

.....

*Cet article traite un aspect de l'organisation de l'allocation des tours de parole dans l'interaction institutionnelle: l'emploi des prénoms et leur démarcation prosodique dans la sélection du locuteur suivant. L'interaction institutionnelle se caractérise par des droits asymétriques à prendre la parole ainsi que par la préallocation de l'action. Cela implique la restriction d'une partie à poser des questions et de l'autre à y répondre. Les analyses se centrent sur deux types de situations formelles multipartites: une situation de salle de classe en coprésence et une émission télévisée interactive en direct avec des participants à distance. Dans chaque contexte, l'allocation des tours est déterminée par une partie: l'enseignant ou l'animateur de l'émission télévisée, qui, après avoir posé une question, désigne le locuteur suivant par son prénom. L'emploi des prénoms est examiné ici d'une manière locale en termes d'organisation des tours et d'attention mutuelle vers la sélection du prochain locuteur: comment la démarcation prosodique est-elle sensible au contexte pour ce qui est de l'accès visuel et de quelle façon le prénom est-il employé pour attirer l'attention? Les descriptions détaillées de l'ordre séquentiel sont basées sur l'analyse*

*conversationnelle d'inspiration ethnométhodologique. L'objectif consiste à comparer quatre cas précis, tout en préservant les traits spécifiques de chaque extrait et l'essentiel de la situation.*

**Mots clés :** *tour de parole, analyse conversationnelle, prosodie, situations multipartites, linguistique interactionnelle, salle de classe, interaction de médias*

## 1. Introduction

1 It is difficult to give a full definition of institutional talk or to define exactly the boundaries between “conversation” and other types of talk, which have some distinguishable aspects. Six dimensions of difference have been systematized (see Drew & Heritage, 1992): i) turn-taking organization, ii) overall structural organization of the interaction, iii) sequence organization, iv) turn design, v) lexical or word choice, vi) epistemological and other forms of asymmetry. Our analysis will focus on the first one: turn-taking organization. For conversation analysis, understanding turn-taking for conversation and other forms of talk-in-interaction is key to understanding human conduct. All interactions involve the use of some kind of turn-taking organization (Sacks et al., 1974) and many types of institutional interactions use the same turn-taking organization as ordinary conversation, considered as a fundamental domain of social interaction (Heritage & Clayman, 2010: 12). Sacks (1992) laid out a basic systematics for turn-taking in conversation by comparing it to turn-taking in other forms of talk-in-interaction.

2 In institutional interactions, topics, actions and turn allocation are organized in a more predictable way. These interactions, where participants carry out tasks related to their institutionally relevant identities, involve some very specific and systematic transformations in the conversational procedures of turn-taking. These modifications in the turn-taking system, characteristic of institutional interaction, have the potential to alter participants’ opportunities for action. In effect, turn-taking practices organize the allocation of opportunities to participate in conversation and the turn-constructive forms such participation takes. According to Heritage (1998: 293), a specific turn-taking organization is necessary in formal environments where there are many potential participants in the interaction, whose contributions must be formally distributed, and/or when the conversation is designed for a “listening” audience. The most pervasive form of pre-allocation of action involves the restriction of one party to asking questions and the other to responding to them (Heritage & Clayman, 2010). Using a question as a sequence-initiating action restricts who can speak and when, and the type of turn that can be produced. This specific turn-taking organization, where turn allocation is determined by one party, shapes participation within sequences of action. Most actions performed through talking are shaped by the organization of that talk into speaking turns: how speakers compose their contributions, where they position those contributions in the ongoing interaction, when they get to participate (Lerner, 2004: 4).

3 In ordinary multiparty conversations, a question may be answered, either by anyone present or by the next speaker who has been explicitly selected. In formal interactions,

the action sequence operates *dependently* of selection of the next speaker. Therefore, selecting someone to speak next is accomplished by combining two elements in the turn-at-talk: i) a first pair part, and ii) some form of address (Lerner, 2003). This explicit allocation of turns characterizes interactions in courtrooms, news interviews, and classrooms (Heritage & Clayman, 2010). In these two situations, selection of the next speaker is done by using the first name of the co-participant.

- 4 The data we analyze in this paper are part of two of these categories of institutional settings characterized by action pre-allocation and asymmetrical rights to talk: on the one hand, the classroom framework where the teacher selects one student among the whole class to answer the question, and on the other hand, interactive media events where the host in the studio distributes turns, selecting a guest from those present in the studio and/or from those who are participating from a distance. Studies in conversation analysis have shown that the most commonly used technique to select the next speaker involves the use of address terms and their position in the turn-at-talk (Sacks, 1992: 665). Sacks makes a clear distinction between choosing someone as next speaker (*addressee*) and using an idea to talk about him (*doing referencing*). Using data recorded in multiparty formal frameworks, the analysis will focus on the situated use of first names to select the next speaker, and examine how he/she situationally becomes the addressee of the question asked by the teacher or host. By this explicit addressing practice, the action making a next turn-at-talk relevant – a question – is contextualized as being addressed to this particular participant. The teacher or the studio host in the extracts analyzed here sequentially organize the course of their actions with the use of the name of the selected participant, either by preposing the term of address as: [first name] [pause] [turn-at-talk], or by postposing it as: [turn-at-talk] [pause/hesitation] [first name]. In the following we will first provide an account of next-speaker selection in institutional settings, and then the paper focuses on the interplay of gaze, first name addressing and its prosodic realization in the two classroom situations we have analyzed. Finally, two extracts from the context of an interactive French television program are used to contrast and discuss how the first name is produced, and where it occurs in the turn selection.

## 2. Next-speaker selection in institutional settings

- 5 The main difference between institutional talk and non-institutional talk concerns the speech exchange system and the organization of turn-taking<sup>1</sup>. Participants in classroom conversations orient to an asymmetric situation where one party – the teacher – asks questions and the other party – the cohort – answers the questions (McHoul, 1978; Mehan, 1985; Atkinson & Drew, 1979; Greatbatch, 1988; Heritage & Roth, 1995; Haddington, 2005; etc.). The communication between teacher and student has its own speech exchange system, which maintains the organization

---

1. See in particular: Drew & Sorjonen, 1997; Heritage & Greatbatch, 1991; Heritage, 1998; Clayman & Heritage, 2002.

“one speaker speaks at a time”. The conversation is achieved in a coordinated way through the organization of the selection of a student by the teacher, so that the addressed participant takes the floor, and no-one but him.

6 In his work on “explicit addressing” for next-speaker selection, Lerner (2003) pointed out that the explicit form of address – addressing a co-participant by name – seems to be deployed selectively, particularly in sequential environments where these forms of explicit addressing do more than just indicate who is being addressed. Indeed, in addition to how an address term is voiced, in particular a person’s name, explicit addressing can be used to accomplish a wide range of actions (Lerner, 2003: 198). For example, in news interviews, terms of address are present in the presentation of misalignment actions: topic changes, non-compliant responses and disagreements (Clayman & Heritage, 2002; Raymond, 2003). They can also serve as resources for the management of certain expressive properties of speech, such as the presentation of points of view as being particularly important or sincere (Clayman, 2010).

7 Our study examines how conversational skills are supported by embodied communication, involving the position of the first name within the sequence-initiating turn, gaze and a specific prosodic marking. In this sense, it is demonstrated that the meaning of the conversation, the syntax and prosody, are affected by the projectable aspects of the turn (Selting, 2000: 512; Goodwin & Goodwin, 1987; Ford & Thompson, 1996). Concerning prosody, the organization of turn-taking appears to be very important. For example, the distinction between “What” as a single word question, and as the beginning of a phrasal sentence construction is not done syntactically but through intonation (Sacks et al., 1974: 721). However, following Schegloff (1996: 116), our concern in this paper is with “the bearing of turn-taking contingencies on certain deployments of prosody, rather than the bearing of prosody on turn-taking organization”. It is thus in terms of the turn-taking organization in the classroom or in TV broadcasts that this paper will be interested in the prosody which accompanies the “explicit addressing” through the use of first names for next-speaker selection.

8 As a method for specifying the recipient to whom the talk is addressed, address terms abound in institutional contexts, which are characterized by a normative activity structure of question/answer and a multiparty participation framework associated to it. These types of institutional contexts – classroom interactions and TV broadcasts, that we analyze in this paper – are favorable for studying phenomena related to terms of address, because the teachers and the studio hosts address the students and the studio guests very frequently. In the analysis of the data collected both in the classroom (between co-present participants) and TV shows (remote and mediated interaction), a prosodic marking seems to accompany explicit addressing in next-speaker selection. The shape of the prosodic marking situationally varies with the context of occurrence. The feature shared by all the various manifestations, however, is that the last syllable of the name always carries a rising tone.

9 These contours appear in co-present situations, when parties are face-to-face and the teacher is gazing at the pupil he is selecting and the latter is gazing back.



The same types of prosodic markings are also used by the hosts in the mediated interaction of TV shows, when the current speaker is looking towards the areas of visualization, i.e., the television screens in the studio where those participating from a distance are visible for those present in the studio.

10 This paper will focus on these two situations, where the teacher and the selected pupil are gazing at each other and where the current speaker is looking towards the areas of visualization in the mediated interaction of TV shows. In these situations, the first name of the selected next speaker is prosodically marked. We will see how the prosodic markings are contextually related to the organization of turns-at-talk in the two formal settings.

11 The analyses aim at answering the following question: how is the prosodic marking relevant, and what is interactionally achieved by it in each situation? The extracts examined here are part of a corpus of data collected in primary and secondary schools in France, and of five episodes of the television show called CULT, broadcast on France 5 in 2005<sup>2</sup>. We have a total of 37 occurrences of address terms. For reasons of space and coherence in this paper, we are going to focus on 4 occurrences that are analyzed in fine detail. The educational data were recorded using four video cameras, professional microphones and a “MiniDisc”<sup>3</sup> audio recorder. In the talk show, four teenage guests participate from their respective home. They are virtually present through video conference and are called “bloggers” in this program. They follow the TV broadcast and all the conversations on their PCs and, when selected by the host, they answer the question, giving their opinion about the current topic. Discussions are mainly about issues related to news or technologies. Each blogger communicates simultaneously with the studio through his/her webcam and microphone at home. During the TV time called the “video chat”, for part of the program, one or two facilitator(s) and one or two celebrity guest(s), who are present in the studio, talk alternately with the four remote participants.

12 This specific situation of audiovisual/video conference connection does not allow any possibility of eye contact or mutual gazing for the coordination of turn-taking between the participants co-present in the studio and those at a distance. Participants look at cameras and/or monitors. Goodwin (1979 and 1980) and others since (Lerner, 2003) have shown how gaze is central to the management of turn-taking. Directing one’s gaze to a co-participant is the first of the two forms of explicit addressing practices. The second one is the practice under analysis here: addressing a co-participant by name or other address term. The co-participant to whom the speaker’s gaze is

---

2. We warmly thank Bruno Bonu, who allowed us to discover the CULT talk show when it was broadcast on television, for his relevant remarks about the contextual use of technical equipment.

3. Ten Have (2007: 79) distinguishes several categories of recordings of “naturally occurring” interaction that can be useful for “doing conversation analysis”. The data of the TV-program CULT are part of the strategy of “copying radio and TV broadcasts” while the classroom data are part of “making one’s own recordings” (another strategy would be to use existing recordings). These video data were collected after having obtained participants’ consent and official authorization.

directed is the recipient, the *addressee*, whose turn-taking is expected according to the *current-selects-next* technique. However, in most technologically mediated communications, like this particular type of visiophony, mutual gazing is not possible<sup>4</sup>.

13 In the same way, though classroom situations are characterized by co-presence, they involve multiple body orientations and commitments of participants. According to Kääntä (2010), in classroom situations, teachers typically allocate turns by using the selected student's name and by directing their gaze towards the student. Teachers may also use embodied turn allocations, such as head nods and pointing gestures that can occur either together with the student's name, with a discourse particle, or on their own. The embodied allocations derive their interactional meaning through their sequential position. However, Kääntä's (2010) results show that embodied allocations are successful only when the participants establish mutual gaze<sup>5</sup>. This, in turn, assigns greater import to aspects of voicing as an interactional resource, because as we will see in the analyses of the extracts, mutual gazing is not always possible<sup>6</sup>.

14 The analysis of the CULT data shows that, in his management of next-speaker selection through the practice of addressing the participant by name, the host uses specific prosodic means. Prosodic marking of the first name thus solves two major practical problems that have to be managed by the hosts of the television program: i) the allocation of the turn to the next speaker, but also ii) the selection of the relevant image set at the right moment, which must be accomplished by the production team. In the case of the interactive TV show program CULT, it is possible for the host to address a blogger participating from a distance and select him/her as the next speaker. While there is a necessary selection by the production team of how (and when) the participants have access to the image of one or the other, they have a constant audio access. It is thus possible to address a distant co-participant by name without actually "seeing" his picture on the screen. We will come back to the technical aspects of the visiophony and the practical aspects of producing a TV program in the next part. The issue at stake is the position of the next-speaker selection within the action-initiating turn. In CULT, the explicit addressing by name is placed before the host's question. In contrast, in classroom interactions, the first name is placed after the teacher's question. Depending on whether the term of address is *pre-positioned* or *post-positioned* (Lerner, 2003), what it achieves is somehow different. We will see how by analyzing first the classroom extracts, and then the CULT extracts.

- 
4. Complex and expensive video conference devices developed by large companies like Orange or Cisco, whose sales argument is the capacity to restore the conditions of face-to-face interaction, especially for remote business meetings, do allow this, though.
  5. Similarly, in dinner conversations, some participants who are visually oriented to their plate and the activity of cutting their food do not see the gaze of the current speaker. It happens that some speak when they have not been designated, thus showing the context-sensitive vulnerability of gaze as a method of explicit addressing (see Lerner, 2003).
  6. For the same reason, the role of prosody is also particularly important, e.g., in radio broadcasts (Lehtinen, 2008).

### 3. Empirical analysis

15 Our empirical analysis is interested in how the initiated sequence, gestures, bodily attitudes and the use of artifacts constitute the situated context to which the form of the first name for accomplishing speaker selection is sensitive. In adjusting the action, what are the practices for selecting the next speaker – term of address, gaze, and prosodic changes? Are they used in combination or is there a preferred use of one or the other, depending on the situation and the conditions of visibility?

16 Consider the following two segments of talk occurring in economics classes in vocational education. Several days of video recordings were made with the same teacher with his students. In this teaching classroom framework, as in many others, a common practice is that students raise their hands when the teacher asks a question; the teacher then selects a student to answer by addressing him by name. Excerpt [1] presents the case in which the teacher selects a student, Mounir, who answers the question. The theme of the course is the European Union; the teacher has already asked which institutions they know. Now, he asks where the European Union headquarters are located.

[1] Excerpt Mounir

1 **teacher** quelles sont [*\*les villes ou les \*pays où- où il y a les-*]  
**what are the \*cities or \*countries where there is the-**

2 **Nora** [ ( inaudible ) ]



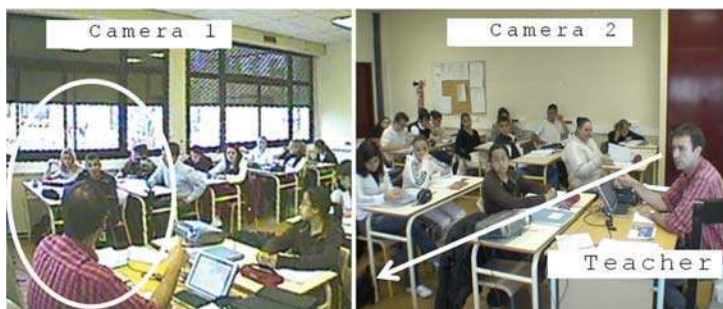
3 **teacher** (0.6) >ça y est< ça y est \*Nora  
**(0.6) are you done, are you done, \*Nora**



- 4 (I.O)  
 5 **Nora** °oui°  
**Yes**  
 6 **teacher** donc (.) \*quelles sont les deux villes  
**so (.) \*what are the two cities**



- 7 euh: se\* euh Mou?ni::↑r  
**uh uh\* Mounir?**



- 8 **Mounir** la Belgique non.  
**Belgium isn't it?**  
 9 **teacher** (.) [oui en Bel?gique,  
**[Yes in Belgium**  
 10 **Mounir** [euh >plus précisément< à Bruxelles:,  
**[uh more precisely in Brussels**  
 11 **teacher** voilà?  
**that's it**

17 During the production of the question by the teacher (line 6), Mounir raises his hand. This action is common at this sequential moment, as studies of teaching situations in ethnomethodology and conversation analysis have shown (McHoul, 1978: 201; Mondada, 2009a: 31). By raising his hand, the student is making himself available for selection which will allow the teacher to categorize those who are able or who wish to answer.

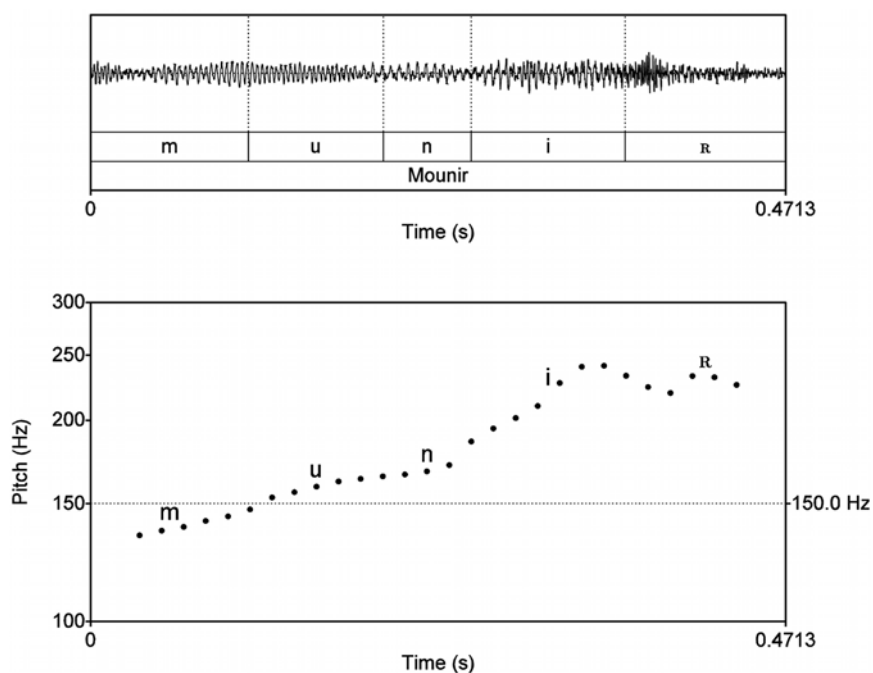
18 The actual selection is achieved by the teacher, by addressing Mounir by name (line 7). It is only after having been explicitly and publicly selected as next speaker that Mounir takes the floor and answers the question (line 8), lowering his arm.

19 In this excerpt, selecting the next student to answer occurs in a situation where the teacher and the student are co-present in the same classroom environment and are in a face-to-face position. The sequential analysis shows that the sequence-initiating action is accomplished by a turn-at-talk that presents several characteristics. One of them is the position of the first name within the turn and its interactional consequences. What is principally achieved, of course, is the selection of Mounir, who produces an initial answer, and then elaborates on it. However, address terms do more than simply specify *who* the current speaker is addressing, depending on the specific circumstances in which they are used, for example, their positioning.

20 Post-positioned terms of address are regularly employed as a device to demonstrate a particular stance towards a relationship with a recipient, under circumstances where that demonstration is particularly relevant (Lerner, 2003: 185). Lerner has shown that addressing is first achieved by gaze and/or tacit forms of address, and then a name is appended to the sequence-initiating action. A post-positioned address term upholds the (already adequately established) intended recipient. It “personalizes” the question for Mounir, who has pre-selected himself by raising his hand, and who is already the addressed recipient, since his reciprocity has been established by gaze (line 6). The explicit addressing by name that the teacher produces also makes it publicly visible, to the other-than-addressed participants, who the selected one is. According to Détrie (2007: 11) and Isosävi (2010: 129), post-positioned terms of address have a jussive value due to the injunction they direct to the interlocutor. In the occurrences we have analyzed, when the first name is placed after the teacher’s question – [question] + [name] –, the post-positioned first name is produced with a continuously rising pitch.

21 The first name “Mounir” [muniʁ] carries a continuous rising contour, the range of which is 10.2 semitones. A pitch rise of this size within two syllables is made possible by lengthening the stressed vowel [i], transcribed as [Mou?ni::↑r]. Graph 1 presents the waveform and the pitch curve obtained with the speech analysis program Praat 5.3.16 (Boersma & Weenink, 2012) during the production of the first name “Mounir”. The finely dotted line indicates the mean pitch level of the speaker. Thus, as can be seen in the graph, the rise starts a bit below the mean pitch level of the speaker and moves towards the upper end of the speaker’s range. According to Mertens’ (2008) “predictive theory” on French intonation, this type of major interval from down to up (LH)<sup>7</sup> invites the interlocutor to react. Morel and Danon-Boileau (1998) presented similar findings.

7. The abbreviations refer to French intonation contours as they are defined in Mertens’ (2008) “predictive theory” of intonation. “L” stands for low and “H” stands for high.



Graph 1. Waveform and pitch curve during the production of “Mounir”

22 McHoul (1978) showed the preponderance of post-positioned first names for classroom interactions. He explained this by the fact that, for the students, what is at stake is their responsibility, as hearers, to listen to any question-in-production for its being possibly addressed to them (McHoul, 1978: 207). Similarly, Lerner (1995) analyzed the post-position of the first name as allowing a delay in the identification of the student who is being addressed. This turn-taking design allows the turn to be treated by any potential participant, until the explicit addressing by first name, which identifies a particular addressee and possible next speaker.

23 Moreover, as we have seen in the excerpt above, mutual gazing may be involved in turn allocation. It is involved in the teacher’s selection of the next speaker in extract [1]. Selecting Mounir is achieved through both gaze-directional addressing and addressing by first name. The video analysis of this face-to-face interaction gives insight into how the combination of the two practices is employed by the teacher: he turns to Mounir (line 2) as he finishes the question’s turn-constructive unit (TCU), which is understood and oriented to by the students – especially Mounir – as a question. When the teacher begins uttering [Mou?ni::↑r], they are mutually looking at each other. In this face-to-face classroom interaction, participants mutually see each other and see each other’s actions. Selecting Mounir is no accident; it has been made conditionally relevant. The teacher produced the second TCU of his turn (“Mounir”) after seeing Mounir raise his hand. By

pre-selecting himself, Mounir is visibly and understandably asking to be selected. As the teacher maintains his gaze on Mounir, he produces the latter's first name with a continuously rising pitch (graph 1).

24 Lerner (2003: 196) demonstrated that several explicit addressing practices can be employed simultaneously/alternately to achieve a sequence-initiating action, though in practice, each of these methods – gaze-directional addressing and addressing a co-participant by first name or other term of address – seems to have some limitations specific to the context. This empirical analysis shows a third element associated with the method of addressing the next speaker in formal talk and face-to-face classroom situations. Indeed, in addition to verbal and embodied means of turn allocation, the teacher employs a third interactional resource: the prosodic marking on the first name.

25 Excerpt [2] presents another instance in which the same teacher confirms the giving of the floor to another pupil, Maryam. The teacher asks students to comment on economic news. He asks what the amount of tax based on salary is.

[2] Excerpt Maryam

1 **teacher** tout le monde.\* est d'accord,  
**does everyone\* agree**



2 **Maryam** non c'est pas plu:s\* c'est égal que::  
**no it's not plus,\* it's equal to**



3 **teacher** ou:i (.) Ma?\*ryam?  
**Yes (.) Ma\*ryam**



4 **Maryam** que son salai\*r:e (o.5) c'est sur les heures qui vont mettre en  
**to his sala\*ry, it's on the overtime hours**



5 trop elles s'ront pas euh:

**they will not**

6 **teacher** elles s'ront pas,

**they will not**

7 **Maryam** elles seront pas taxées

**they will not be taxed**

8 **teacher** voilà c'est bien donc elles seront pas taxées

**here it is good then they will not be taxed**

26 Maryam, who had begun to express her disagreement in response to the teacher's question (line 1) interrupts her turn with a hesitation. She only takes the floor to continue her answer (line 4) after the tacit form of selection by the reciprocity token "yes" (line 3), followed by the explicit addressing produced by the teacher (line 3). We know that the class event is under the teacher's responsibility (McHoul, 1978). Thus, the turn-taking organization – management and turn allocation – and the privileged access of the teacher to the floor, involve the maintaining of shared attention among the students as a practical problem. Here, the students' attention is organized around the question-answer sequence and the turn allocation by the teacher, who has specific rights: i) to initiate actions by asking the questions, and ii) to distribute turns. Maryam orients to this specific turn-taking organization by starting an initial answer to the question (line 1) addressed to "everyone". It is important to note that the teacher looks at Maryam only as early as the beginning of his reciprocity token "yes" (line 3). Although the occurrences of "yes" associated with



the exchange of glances are recognizable and predictable units of talk (Sacks et al., 1974), Maryam does not continue to talk after the teacher's assessment. Maryam speaks again, completing her answer by taking her turn "where she had left it off" after the teacher has uttered her first name. The turn to select the next speaker Maryam is built, once again, with a term of address accompanied by a prosodic marking (line 4). However, the fact that there is a pause between the reciprocity token "yes" and the first name affects the form of the prosodic marking: the pause creates a prosodic boundary after the word "yes". This leads to the production of an initial accent (IA) (Mertens, 2008) in the beginning of the first name (Morel, 2010; Mertens, 2008). The occurrence of the pause, in turn, is related to the action of the participants, i.e., to the fact that Maryam does not continue talking after the "yes", but she waits until she hears the teacher utter her name.

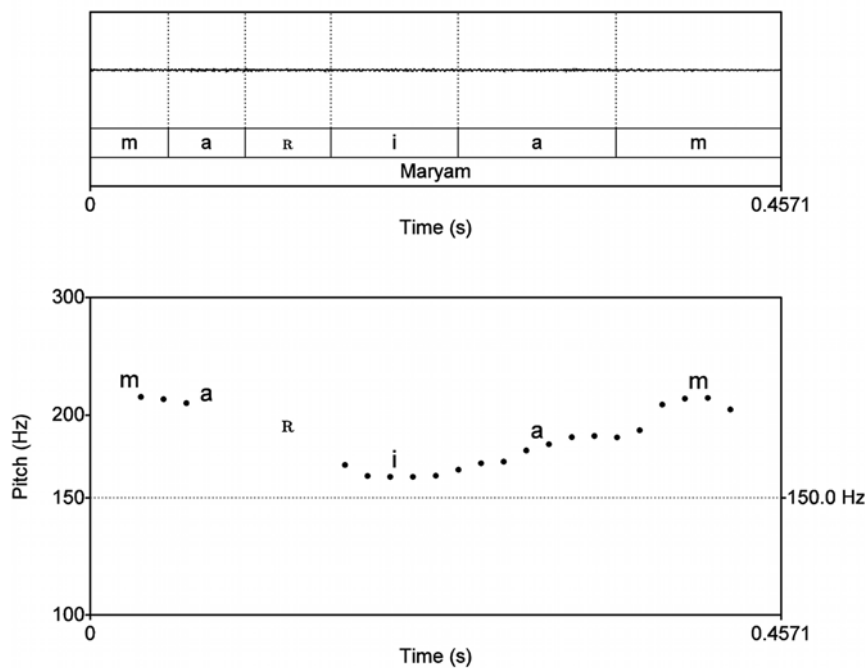
27 In this excerpt, the explicit address to Maryam was preceded by "yes" and by the gaze in her direction. We can observe that the prosodic marking on the first name optimizes the next-speaker selection, by making speaker change clear and identifiable for all the participants, including those who have not been addressed. In this occurrence, the teacher uses a turn-taking organization format as a means to intelligibly and explicitly manage the progression of the participation.

28 Due to the fact that the teacher's turn includes two consecutive actions – [reciprocity token] + [term of address] – divided into two separate prosodic groups, the form of the prosodic marking of the name is rather different from the one in excerpt [1], where the name was marked with a continuous pitch rise (LH) from the beginning until the end. Indeed, the first syllable [ma] of the first name carries an IA that appears in the graph as a raised pitch level (211 Hz)<sup>8</sup>. During the production of the unstressed penultimate syllable [ri], there is a pitch fall (4.6 semitones) during which the fundamental frequency (F<sub>0</sub>) approaches the mean pitch level of the speaker. After that, the pitch level goes up again during the production of the last syllable [am]. The size of the final pitch rise that takes place here is only 4.8 semitones, which is rather modest. In Mertens' (2008) predictive theory of French intonation, this rise would be at the limit of a minor and a major interval<sup>9</sup>, whereas the one occurring in the first example [Mou?ni::↑r] clearly constitutes a major interval. We have noticed throughout the data examined in this paper, that final rises marked with minor intervals occur in face-to-face classroom arrangements, whereas in the two excerpts of TV broadcasts, the final pitch rises are larger. However, a movement of this size is still sufficient to create the effect of a rising tone. A pitch contour of this type, starting with an IA, works to emphasize an entity and/or to start a new information unit (Mertens, 2008), since where a rise of the type LH links the word prosodically to the preceding group, a word starting

8. In addition to a raised pitch level, the IA of French also typically entails an increased level of loudness, a strengthened phonatory force and a longer duration of the syllable (Di Cristo, 1998; Mertens, 2008).

9. The size of a minor interval is less than 4-5 semitones, depending on the size of intervals used by the speaker.

with an IA starts a new prosodic group (Mertens, 2008; Wichmann, 2000). In this context, the IA emphasizes the term of address. The need to emphasize it with the help of the IA comes from the fact that Maryam has not interpreted the reciprocity token “yes” as a sign of turn allocation. Prosodically, the IA constitutes a marked feature that reduces the importance of the final rise as a marking device. Thus, in this case, the prosodic marking consists of two parts: IA and final rise.



Graph 2. Waveform and pitch curve during the production of “Maryam”

29 Although the size of the interval is different in the two excerpts examined so far (larger for [Mou?ni::↑r] than for [Ma?\*ryam?]), in both extracts, the prosodic marking with a final rise on the first name occurs when the selected student is facing the teacher and both are mutually gazing at each other when the address term is produced.

30 The next section will examine the prosodic phenomenon associated with the first name, using the data of the interactive television show. While in the classroom situations examined above, the first name is post-positioned, that is, it comes after the teacher’s questions, the first name is pre-positioned to the host’s question in the following excerpts. We have just seen that the prosodic marking of the first name of the next speaker is observed in the classroom interaction when parties are facing each other. However, despite the differences in i) the positioning of next-speaker selection within the turn (pre- vs. post-positioned), ii) the type of talk and the constraints associated to it (classroom vs. TV), and iii) the interactional framework

(face-to-face vs. remote and mediated), we observe a similar prosodic marking of the name of the selected next speaker in both situations.

#### 4. The first name in an interactive French television program

31 Before analyzing the sequential occurrence of the rising pitch, it is necessary to describe the physical and ecological media setting from an interactional point of view. Only the hosts and the celebrity guests are physically present in the studio. The teenage “bloggers” participate in the video chat conversation “from a distance”, connected audio-visually via their webcam. In addition to the virtual co-presence of the bloggers (Fornel, 1994), which implies a specific management of the mediated communication, the technologized communication with the production-control room (Broth, 2008 and 2009) occurs simultaneously to the interaction in the TV studio. This means that the host has to manage several parallel mediated interactions.

32 The coordination between the control room and the studio is absolutely necessary for the production of the TV program, and is achieved in several ways. The staff of the control room do not have direct access to the studio space; they carry out their professional tasks by seeing through screens (Licoppe & Dumoulin, 2007; Broth, 2009: 2014) and listening to the interactions taking place in the studio. These studio interactions constitute the media product that the viewers will access in real time. But the live talk is also a “talk at work” (Drew & Heritage, 1992) produced in real time, which is used by the production-control room personnel to broadcast live. We will now see how.

33 In excerpt [3], the topic of the discussion between the hosts and guests is the disaster in Asia caused by the tsunami. Patrick, one of the participants who is taking part via his webcam and mobile phone, has just finished his turn-at-talk (line 1), in which he said that he was struck by the fact that we cannot do anything against nature. Then, the studio host, Chakib, turns to another blogger, Amélie (lines 3-5), who is virtually present.

[3] Excerpt Amélie

1 Patrick la \*natu?re (.) .h \*hh  
\*nature (.) .h \*hh



2 (I.8)  
3 **Chakib** ouais, .hhh A↑médie↑Δ  
**ok, Amélie**  
4 (I.o)\*



5 **Chakib** \*>Amél[ie<=



6 **Amélie** [oui?  
yes

7 **Chakib** =[même question qu'] est-ce qui t'a l'plus frappée?  
**same question what struck you most**

8 **Amélie** [ bonjour Chaki- ]  
**hello Chaki-**

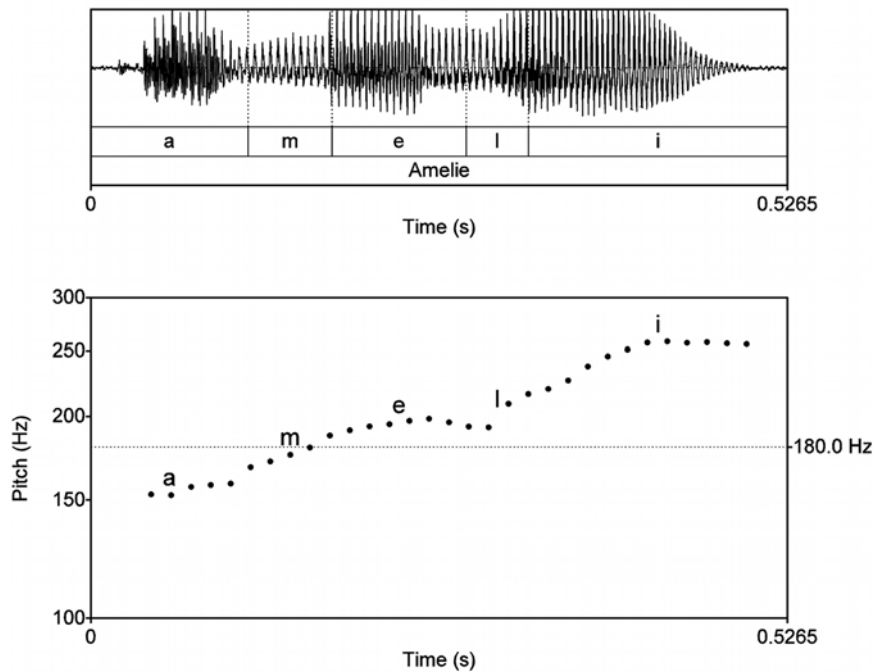
9 \*(I.o)



10 **Amélie** .hh moi c'qui m'a le plus frappée ça c'est euh: plutôt  
**for me what struck me most it's emm rather**

11 l'indécence de certaines euh:: certaines équipes-euh  
**the indecency of some er some teams**

34 The host acknowledges the end of Patrick’s turn-at-talk and produces an assessment, line 3, which serves to close the question-answer sequence previously engaged with him (not transcribed). After a long in-breath, the host selects the next speaker, achieved by the explicit addressing by first name: “Amélie” (line 3). There is a continuous rising pitch associated with an increased level of volume. The size of the pitch rise is 9.3 semitones (line 3), and it is carried by all three syllables of the name. Thus, the rising pitch movement occurring here is similar to the one in excerpt [1], [Mouʔni::↑r]: in both cases, the size of the pitch movement is a major interval from low tone to high tone (LH), and the size of the rise is more than 9 semitones. As already mentioned, according to theories on French intonation, this type of contour (LH) typically invites the interlocutor to react (Mertens, 2008; Morel & Danon-Boileau, 1998). Raising the pitch in a quasi-interrogative fashion may accompany summons items, which is not necessary when the summons stands alone as in “Jim” (Schegloff, 1968: 1081).



Graph 3. Waveform and pitch curve during the production of “Amélie”

35 Here, the term of address “Amélie” (line 3) achieves a summons projecting an answer. This addressing constitutes a conversational action with a prospective orientation. By inviting the next speaker who is virtually present to take the floor, the host orients to the institutional context. The broadcast setting is characterized by a formalized speech exchange system, with turn-type distribution rules and pre-allocation of turns. By doing next-speaker selection and topicalizing the fact that Amélie has been selected to answer the “same question” (line 7), the host acts according to the

restrictions associated with this type of media talk. In addition, as shown on the graph above, the host modulates his voice when selecting the next speaker.

- 36 Amélie’s answer, which has been projected by her explicit selection, is not produced immediately. After a one-second pause (line 4), the host repeats the first name (line 5). In the situation where the called party is located at a considerable distance from the caller, Schegloff (2007: 18 and 52) shows that when the called party does not respond quickly to his being called by his first name, the caller repeats the call with a different kind of prosody. In the excerpt analyzed here, the host orients to the lack of taking the turn, which is projectably expected, by the next speaker. The “no response” from the selected speaker shows how managing transitions in talk is a practical problem. In this sense, pre-positioned terms of address are regularly employed as a device to establish or verify the availability of a recipient in situations where this may be problematic (Lerner, 2003: 184). Indeed, pre-positioned terms of address have a calling value in French (Détrie, 2007: 111; Isosävi, 2010: 129).
- 37 The answer of the addressed recipient is produced line 6, “yes”, showing that she understands the action of the host as being a summons. The summons/answer pair in the form of first name/acceptance occurs as a pre-sequence (Schegloff, 2007: 52) for the exchange of opinion on the topic of the tsunami in Asia. The host is now asking Amélie the same question he had asked Patrick and explicitly signals this topic continuity (line 7). Contrary to the two previous extracts from classroom interactions – where the terms of address are placed after the question, as anyone is a potential answerer till one pupil is selected –, pre-positioned address terms specify the addressee before the question is asked. The analysis of this pre-sequence shows that the pre-allocation of turns by the host, in this mediated communication situation, takes a particular form by which the host ensures the availability and attention of the specifically selected remote participant. By the slight delay in Amélie’s response, the technical aspect of the communication appears as a practical problem, which has to be managed in this context of mediated interaction. Therefore, the host’s summons acts to: i) select Amélie as the next participant in an accountable way, ii) verify whether the selected participant can hear the question and understand that it is addressed to her, but also iii) solve the practical problem of coordination with the production-control room personnel.
- 38 The professional activity of selecting images of the bloggers/co-present participants during the production process, reflexively makes this potentially problematic aspect of speaker transition visible and analyzable, in this doubly mediated – remote and media – interaction. Based on the detailed transcript of extract [3], we can see that the close-up shot of the host appears after a one-second pause (line 4). The control room team shows the current speaker in action on the screen. Since the expected turn-taking by the selected speaker is absent (line 4), the change of framing will continue on a *split-screen*<sup>10</sup> where all the four remote participants (bloggers) are

10. See Mondada (2009b).

visible (line 5). Indeed, the practice of realization is that during the work of putting in images, the team may choose to send on air the image of the current speaker or, alternatively, a listener (Broth, 2008: 69). At this sequential moment of speaker transition, the split-screen leaves open the possibilities for the effective turn-taking of the selected speaker. We note that the image of Amélie is just one among those of the four participants who are potential speakers (image highlighted by the white circle). It is only when Amélie becomes the current speaker by producing the greeting, and the interaction is properly opened – the greeting was preceded by the summons/answer sequence –, that the camera focuses on her (line 9). By choosing to screen the image of Amélie, the control room team make their understanding of the practical problem of turn-taking visible: they have been following the interaction in order to be able to produce the action of selecting the right image at the right moment. The summons “Amélie” is also audibly available to them, and – given their next action of producing the close-up shot of Amélie – they mobilize it as a resource for their professional activity. The occurrence of “Amélie”, which is explicitly addressed to her as part of the media talk, is also mobilized as a “talk at work” for the coordination between the host and control room team for the production of this TV program.

39 We know, based on the work of Macbeth (1999), that there is a reflexive relationship between the practice of filming and what is actually being filmed. In this section, analyzing what is achieved in the two excerpts [3] and [4] shows that the shot by the control room and the organization of speakership are reflexively embedded. First, a special coordination occurs between the host(s) and the control room for the production of this particular TV program with distant bloggers. Then, there is a spatial division of the stage where the co-present participants (the hosts on one side and the invited guests on the other) are seated facing the other group (line 1). This ecological organization of the TV show, where the four bloggers are visible on screens attached to walls, allows the cameras to be placed on predetermined axes (Relieu, 1999), without the need for active shooting: the video recording cameras are arranged on a circular rail around the participants and are controlled from a distance by the production-control room. Thus, the control room team directly selects the shots to be broadcast live without the intervention of cameramen, and without the possibility of the hosts giving them information. The live TV product is the data to which the present study has access and which allows a reflexive analysis of the real-time editing work that is achieved.

40 By these complex technological processes, the live TV broadcast is the fulfillment of a “connected co-presence” (Licoppe, 2004): the bloggers are seated at their personal home computer, and a telecommunications connection remains open for the vocal, verbal, and visual resources. The participants take part at a distance by making gestures, showing objects, their room environment, etc. In the studio, turn-taking is organized, either in relation to what participants have to say about the current topics, as we saw in excerpt [3], or in relation to remote participants’ actions. They can express their interest in the topic at hand, through gestures and “showing practices”. This is an efficient way to manage “mediated” interactions, by showing active participation

without having to speak while one has not been selected to do so. Though turns are allocated by the host in a formatted and restrictive way, turn allocation is not necessarily determined in advance. Its local production is achieved by following the conduct of both the guests present in the studio and the bloggers. Extract [4] is an arrangement of this type. In this situation, the studio hosts are talking with the co-present guests about the soundtrack of a video game. In line 8, one of the two hosts, Ray, suddenly calls one of the distant participants, Patrick, by name, and takes the floor over the other host, Chakib, who was initiating his turn at the same moment (line 7).

[4] Excerpt Patrick

1 Maud euh\*:: a:, [débuté-  
er\* has started



2 Ray [>comb-<  
how ma-

3 F [ha?  
eh

4 Ray com?bien: de cédé?s:  
how many CDs

5 Maud il y en a huit?t: (0,5) hu:it \*comme les huit radios?,  
there are eight (0,5) eight \*like the eight radio stations



6 (.)

7 Chakib [y a-  
there are

8 Ray [Pa↑trick↑Δ



- 9 Patrick ((reprend la boîte)  
 ((takes the box again))  
 10 Chakib ouai:s, Patri:ck.  
 yeah Patrick  
 11 Ray >tu\*asl-< tu l'as déjà?::Δ  
 do you have it already



- 12 (0.7)  
 13 Patrick ah?:: c'est mon jeu ça?:  
 ah that's my game  
 14 Ray ben: tchu as £le dje?:u: [wou\*?:i,  
 well you have the game yes  
 15 Maud [£he\*hh



- 16 Ray tu joues \*combien-  
 you play \*how many-



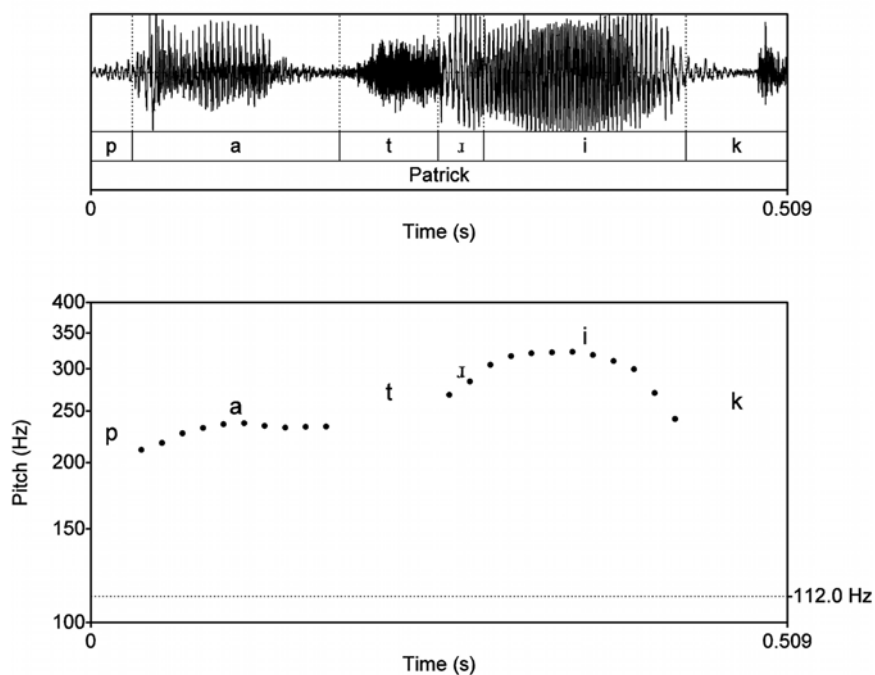
41 In line 2, the host asks a question (line 4) which is immediately followed by an answer from Maud, a columnist. The question continues the sequence, which concerns the content of the video game box. Yet, while the current speakers are the co-present participants, the production team has chosen to show the blogger Patrick (line 5), since at the same time as the question-answer sequence between the host and the columnist is produced, Patrick waves in front of his webcam, the video game box (line 1) that is being referred to in the conversation.

42 Following the micropause (line 6), the two hosts simultaneously take the turn, which results in an overlap (lines 7 and 8). Ray summons Patrick by addressing the latter by name, while Chakib starts referring to Patrick, interrupts his turn and confirms the selection “yeah Patrick” (line 10). By initiating their turn at the very same moment, the hosts show their alignment as to the understanding of what is being displayed – on their control screen and for the viewers – by the control room. They orient to it in the same way. By their simultaneous selection of Patrick, the two hosts make it visible that they have seen the image put out by the control room, that they have understood the relevance of this selection, as well as Patrick’s action. Patrick achieves participation in the interaction by his showing gestures, even if his participation is not verbal. It serves as a pre-selection to speakership, which is confirmed by the hosts’ next action, which consists in allocating the turn to Patrick.

43 The occurrence of the summons “Patrick” (line 8) is characterized by striking prosodic changes. In addition to the large pitch rise (7.4 semitones), the term of address is marked with a strong increase in volume. In allocating the turn (line 8), Ray is perceptibly shouting. These strong prosodic changes may be explained by the turn-competitive situation, since the two hosts – Ray and Chakib – initiate their turns at the same time (Kurtić, Brown & Wells, 2009; French & Local, 1983). It is also noticeable that the last syllable [ɹɪk] of the name carries two consecutive pitch movements: a rise (that starts already during the first syllable [pɑt]) and a fall. The fall is also relatively large: 5.0 semitones. It is, nevertheless, 2.4 semitones smaller than the preceding pitch rise. The duration of the fall is also remarkably shorter than that of the rise: the rising part of the name is 0.35 seconds long, whereas the duration of the fall is only 0.16 seconds. In addition to this, there is a decrease in intensity (-3.6 dB) during the falling part as compared with the rise. Because of these acoustic features, the fall is less prominent than the rise. Consequently, the pitch rise can be considered to be the main constituent of this prosodic marking. Graph 4 below presents the waveform and the pitch curve during the production of this first name.

44 According to intonation theories of French, this type of contour (HL) – where the final accent of an intonation group consists of a pitch rise followed by a pitch fall – typically works to “capture the attention of a person in a kindly manner” (Di Cristo, 1998: 216) or to focus the attention of the interlocutor on something (Mertens, 2008). In this example, it seems obvious that Ray’s turn (line 8) acts not only as a selection of the next speaker, by establishing Patrick as the addressee of the

forthcoming question, line 11, but also as a summons of a specific attention-capturer kind. Indeed, it is frequent in interaction that one linguistic structure has several simultaneous illocutionary values (Kerbrat-Orecchioni, 2001: 34). The “double role” of the turn may explain here the use of this particular type of prosodic marking. According to studies on French phonetics, this type of intonation contour may also include an implication of the speaker involvement (Delattre, 1966; Portes, 2004), and can be used to express emotions, such as joy or disappointment (Léon, 1971: 54). In this example, the rising-falling movement seems to emphasize the speaker’s enthusiasm.



Graph 4. Waveform and pitch curve during the production of “Patrick”

- 45 This prosodic marking of the turn, which achieves next speaker selection (line 8), optimizes the giving of the floor to a remote participant by making the selection clearer. The host uses a summons to establish the attention of the remote participant (Schegloff, 2007: 50; Isosävi, 2010: 130) while the latter is already on screen. Schegloff (1968: 1081) has shown that the term of address as part of the summons may include a rise in the tone of voice. In addition, this seems to be generally the case when the summons comes after the sequence that has already been initiated, as in extract [3]. However, the rising tone may not be necessary when the first name is the only element composing the summons. In this occurrence, the term of address “Patrick” (line 8) is used alone. Nevertheless, the increase in pitch and volume is very marked. It is likely that some differences, compared to the phenomena described

by Schegloff, are due to situational differences. The analysis of the occurrence of “Patrick”, including this specific prosodic marking, reveals the technical specificity of this remotely mediated interaction. There is no direct visual access between the participants, who interact through the device and see each other through the mediation of screens: computer screens for the bloggers at home, or monitor screens for hosts present in the studio. It appears that “seeing through screens” allows for an intelligible understanding of the ongoing activity and an appropriate participation in it, for distant participants like Patrick, as well as for co-present ones.

46 In line 5, the blogger, who is facing his computer screen and webcam camera, is seen checking on his screen whether his showing the CD case is visible, in the same way as it can be seen in the live broadcast. Patrick therefore adapts his conduct to the specific context of mediated and broadcasted interaction: i) his showing practice takes into account the mediation possibilities, as well as the constraints of a restricted angle shot, through the webcam, and ii) as already mentioned, he waits till he is explicitly selected and addressed to participate verbally. When the attention-capturer summons is produced, Patrick is already engaged in the ongoing conversation (line 1), showing his commitment and understanding of the current topic by his action of showing the CD case being discussed. His actions in front of the webcam reflexively contribute to the meaning of the scene (Heritage, 1984: 104). In the same way, the situated selection of the next speaker, which is our object of study, reflexively relies on and achieves this mutual intelligibility.

47 Like Patrick, who fulfills his role as an active distant participant in the show, the host fulfills his role of host. He stimulates discussion and distributes turns by addressing Patrick and explicitly topicalizing what the viewers can already see. Contrary to the summons “Amélie” (excerpt [3]) produced by the host, and relevantly used as a resource by the production team, for Patrick, the production team’s action of selecting one participant’s image *precedes* the hosts’ selection of the next speaker. Next-speaker selection by the hosts is therefore conditionally projected as the next relevant action. Patrick’s action accomplishes relevant and appropriate participation: he shows commitment but does not interrupt the ongoing talk. Pre-selection is achieved and is oriented to by the control production team. Following Patrick’s action – which is selected as the relevant image by the production team and made visible through the mediation device – the hosts orient to giving the floor to Patrick. Thus, it appears that “the selection of the next speaker” is a collaborative in situ achievement. It is achieved by: i) the two hosts, mainly by Ray who addresses Patrick – the selection which we have been analyzing –, ii) Patrick’s action, whose participation is interpreted as a pre-selection, and iii) the control team who, by selecting Patrick’s image, choose to draw the hosts’ and the viewers’ attention to Patrick’s action.

48 As an answer to the summons addressed to him (line 8), Patrick again places the CD case in front of the webcam (line 9), while the guests watch the screens to see what he is doing (line 15). Given the co-present participants’ orientation – Ray

addresses his question to Patrick and the other participants are seen laughing or smiling –, this action appears to be the appropriate response to the summons (Schegloff, 1968: 1080). Patrick becomes the addressee and the CD case he is showing, the object being talked about. The shared attention of all the participants – be they co-present or remote – is achieved. Patrick’s participation therefore raises an important practical issue, which will lead to our final discussion about visualization and next-speaker selection.

49 Watching the remote bloggers through the several screens available in the studio constitutes a practical problem for the hosts but also for the guests. We know that “multiparty communications [video conferencing] tend to be organized in such a way that all the participants are visible, especially the one who is speaking” (Fornel, 1994: 116 [our translation]). In the situation analyzed here, the simultaneous images of the bloggers are arranged at different places in the studio. Line 15 shows that Maud, the columnist, who is seated on the left of the bench, is watching the screen placed in front of her when she laughs. The body torque (Schegloff, 1998) – an intermediary body position relative to the main position – of the other two participants sitting on the bench (line 15) allows us to observe that they have turned to look at the same screen as the host who is talking to Patrick. The body torque of these three participants shows that their attention is oriented towards the action that is being shown on the screen. Contrary to face-to-face situations, in this studio mediated interaction, a participant cannot adapt to the physical movement of the current speaker by changing his own positioning and moving his field of vision. He is dependent upon what the screen offers (Fornel, 1994: 112). The fact that the interaction takes place within this particular framework – technology-mediated media talk – demonstrably influences its arrangement and content. The context or framework of the live video chat within which the talk occurs is relevant to the parties: they orient to the situation and adjust their “way of seeing” by changing their body positions to watch the screen. It thus appears that the forms of interaction that are produced, including the terms of address and their specific prosodic marking for next-speaker selection, are context-sensitive in a “procedurally consequential” (Schegloff, 1991: 49) way.

## 5. Conclusion

50 Next-speaker selection is achieved, in all four extracts analyzed here, with the help of a prosodic marking carried by the first name. The shape of the prosodic marking varies with the context of occurrence, with three prosodic contours. The common feature of all these different manifestations is that the last syllable of the name carries a rising tone.

51 The rising tone occurs regardless of the position of the term of address and what it achieves within the ongoing talk. In the classroom data, the question is uttered before the designation of the speaker selected to answer. In this way, the question

is addressed in a general way to the cohort and any student present in the classroom is a recipient of the question. This post-positioning of the term of address in the classroom situation contrasts with the summons, which is produced as a pre-sequence to the “main” question-answer sequence, characteristic of media talk. Analysis has shown how the post-positioning of student selection achieved joint attention to the classroom event, and how the pre-positioning summons succeeded in attracting the attention of the selected guest in the distant-media-mediated interaction. Apart from the type of institutional talk, the way joint attention or attention-drawing is achieved reflexively draws on the interactional and praxeological resources that are available to participants in each situation.

52 The rising tone occurs, regardless also, of whether the situation is face-to-face or mediated and varies with context-specific factors. In the classroom arrangement that we have examined, the parties have direct visual access to each other. The exchange of gaze is absent in this public media talk where the mediation is twofold: i) the technological mediation of the “remote” communication by video conference, but also ii) the involvement of the control room team in selecting the images. By putting the image of Patrick on screen, in excerpt [4], the control room allowed the hosts to see Patrick’s action on their monitors, to understand thanks to the team’s selection that Patrick’s action was potentially relevant, and thereby projected the hosts’ action of selecting Patrick as next speaker. Likewise, the host’s selection of Amélie by explicit address, in excerpt [3], allowed the relevant image of Amélie to be put on screen by the control room. Achieving *connected co-presence* (Licoppe, 2004) draws on a complex system of seeing other participants’ actions.

53 Next-speaker selection is a coordinated collaborative achievement in classroom situations also. In excerpts [1] and [2], Mounir and Maryam had pre-selected themselves but in different ways, thus projecting their selection by the teacher. The latter does not choose a student at random from the cohort, but takes into account the contingent actions produced by the pre-selected students. In each extract, the teacher selects as next speaker the student who has shown a visible and understandable interest in answering the question. The selection of the next speaker in the occurrences analyzed here therefore emerges as the result of a reflexively constructed collaborative work.

### Transcript conventions

54 The data were transcribed according to conventions developed mainly by Gail Jefferson and commonly used in Conversation Analysis.

[	overlapping talk
=	latching
(.)	micro pause
(3.0)	pauses or silences in seconds
.hhh	audible out breath

:	extension of the sound or the syllable it follows
.	stopping fall in tone
,	continuing intonation
?	moderate rising inflection
↑	rising inflection
Δ	rising volume
*	shows the point at which the image extracted from the video corresponds to the transcribed word
-	interruption
> <	quicker than surrounding talk

## References

- ATKINSON, J.M. & DREW, P. 1979. *Order in Court: The Organisation of Verbal Interaction in Judicial Settings*. London – Basingstoke: Macmillan.
- BOERSMA, P. & WEENINK, D. 2012. Praat: Doing Phonetics by Computer (Version 5.3.16). Computer program available online: <http://www.praat.org>.
- BROTH, M. 2008. The “Listening Shot” as a Collaborative Practice for Categorizing Studio Participants in a Live TV-Production. *Ethnographic Studies* 10: 69-88.
- BROTH, M. 2009. Seeing through Screens, Hearing through Speakers: Managing Distant Studio Space in Television Control Room Interaction. *Journal of Pragmatics* 41 (10): 1998-2016.
- CLAYMAN, S. 2010. Address Terms in the Service of Other Actions: The Case of News Interview Discourse. *Discourse and Communication* 4 (2): 161-183.
- CLAYMAN, S. & HERITAGE, J. 2002. *The News Interview: Journalists and Public Figures on the Air*. Cambridge: Cambridge University Press.
- DELATTRE, P. 1966. Les dix intonations de base du français. *The French Review* 40 (1): 1-14.
- DÉTRIE, C. 2007. *De la non-personne à la personne: l’apostrophe nominale*. Paris: CNRS Éditions.
- DI CRISTO, A. 1998. Intonation in French. In D. HIRST & A. DI CRISTO (eds.), *Intonation Systems: A Survey of Twenty Languages*. Cambridge: Cambridge University Press: 195-218.
- DREW, P. & HERITAGE, J. 1992. *Talk at Work: Interaction in Institutional Settings*. Cambridge: Cambridge University Press.
- DREW, P. & SORJONEN, M.-L. 1997. Institutional Dialogue. In T.A. VAN DIJK (ed.), *Discourse as Social Interaction*. Discourse studies 2. London – Thousand Oaks – New Delhi: Sage Publications: 92-118.
- FORD, C.E. & THOMPSON, S.A. 1996. Interactional Units in Conversation: Syntactic, Intonational, and Pragmatic Resources for the Projection of Turn Completion. In E. OCHS, E.A. SCHEGLOFF & S.A. THOMPSON (eds.), *Interaction and Grammar*. Cambridge: Cambridge University Press: 135-184.
- FORNEL, M. DE 1994. Le cadre interactionnel de l’échange visiophonique. *Réseaux* 12 (64): 107-132.
- FRENCH, P. & LOCAL, J. 1983. Turn-Competitive Incomings. *Journal of Pragmatics* 7 (1): 17-38.

- GOODWIN, C. 1979. The Interactive Construction of a Sentence in Natural Conversation. In G. PSATHAS (ed.), *Everyday Language: Studies in Ethnomethodology*. New York: Irvington Publishers: 97-121.
- GOODWIN, C. 1980. Restarts, Pauses, and the Achievement of a State of Mutual Gaze at Turn-Beginning. *Sociological Inquiry* 50 (3-4): 272-302.
- GOODWIN, C. & GOODWIN, M.H. 1987. Concurrent Operations on Talk: Notes on the Interactive Organization of Assessments. *IPRA Papers in Pragmatics* 1 (1): 1-55.
- GREATBATCH, D. 1988. A Turn-Taking System for British News Interviews. *Language in Society* 17 (3): 401-430.
- HADDINGTON, P. 2005. *The Intersubjectivity of Stance Taking in Talk-in-Interaction*. PhD thesis. University of Oulu.
- HERITAGE, J. 1984. *Garfinkel and Ethnomethodology*. Cambridge: Polity Press.
- HERITAGE, J. 1998. Oh-Prefaced Responses to Inquiry. *Language in Society* 27 (3): 291-334.
- HERITAGE, J. & CLAYMAN, S. 2010. *Talk in Action: Interactions, Identities, and Institutions*. Chichester – Malden: Wiley-Blackwell.
- HERITAGE, J. & GREATBATCH, D. 1991. On the Institutional Character of Institutional Talk: The Case of News Interviews. In D. BODEN & D.H. ZIMMERMAN (eds.), *Talk and Social Structure: Studies in Ethnomethodology and Conversation Analysis*. Cambridge: Polity Press: 93-137.
- HERITAGE, J. & ROTH, A.L. 1995. Grammar and Institution: Questions and Questioning in the Broadcast News Interview. *Research on Language and Social Interaction* 28 (1): 1-60.
- ISOSÄVI, J. 2010. *Les formes d'adresse dans un corpus de films français et leur traduction en finnois*. PhD thesis. University of Helsinki.
- KÄÄNTÄ, L. 2010. *Teacher Turn-Allocation and Repair Practices in Classroom Interaction: A Multisemiotic Perspective*. PhD thesis. University of Jyväskylä.
- KERBRAT-ORECCHIONI, C. 2001. *Les actes de langage dans le discours: théorie et fonctionnement*. Paris: Nathan.
- KURTIĆ, E., BROWN, G. & WELLS, B. 2009. Fundamental Frequency Height as a Resource for the Management of Overlap in Talk-in-Interaction. In D. BARTH-WEINGARTEN, N. DEHÉ & A. WICHMANN (eds.), *Where Prosody Meets Pragmatics*. Bingley: Emerald: 183-203.
- LEHTINEN, M. 2008. *La contextualisation du discours radiophonique par des moyens prosodiques. L'exemple de cinq grands philosophes français du XX<sup>e</sup> siècle*. PhD thesis. University of Helsinki.
- LÉON, P.R. 1971. *Essais de phonostylistique*. Montreal – Paris – Brussels: Didier.
- LERNER, G.H. 1995. Turn Design and the Organization of Participation in Instructional Activities. *Discourse Processes* 19 (1): 111-131.
- LERNER, G.H. 2003. Selecting Next Speaker: The Context-Sensitive Operation of a Context-Free Organization. *Language in Society* 32 (2): 177-201.
- LERNER, G.H. 2004. *Conversation Analysis: Studies from the First Generation*. Amsterdam – Philadelphia: J. Benjamins.



- LICOPPE, C. 2004. "Connected" Presence: The Emergence of a New Repertoire for Managing Social Relationships in a Changing Communication Technoscape. *Environment and Planning D: Society and Space* 22: 135-156.
- LICOPPE, C. & DUMOULIN, L. 2007. L'ouverture des procès à distance par visioconférence. *Réseaux* 5 (144): 103-140.
- MACBETH, D. 1999. Glances, Trances, and Their Relevance for a Visual Sociology. In P.L. JALBERT (ed.), *Media Studies: Ethnomethodological Approaches*. Lanham – New York – Oxford: University Press of America: 135-170.
- McHOUL, A. 1978. The Organizations of Turns at Formal Talk in the Classroom. *Language in Society* 7 (2): 183-213.
- MEHAN, H. 1985. The Structure of Classroom Discourse. In T.A. VAN DIJK (ed.), *Handbook of Discourse Analysis*. London: Academic Press: 119-131.
- MERTENS, P. 2008. Syntaxe, prosodie et structure informationnelle: une approche prédictive pour l'analyse de l'intonation dans le discours. *Travaux de linguistique* 56 (1): 87-124.
- MONDADA, L. 2009a. La production de l'intelligibilité de l'action: une approche multimodale des procédés de sélection des locuteurs dans les interactions en classe. *Études de linguistique appliquée* 1 (153): 25-40.
- MONDADA, L. 2009b. Video Recording Practices and the Reflexive Constitution of the Interactional Order: Some Systematic Uses of the Split-Screen Technique. *Human Studies* 32 (1): 67-99.
- MOREL, M.-A. 2010. Complexité des marqueurs de frontière dans le dialogue oral en français. In N. ANDRIEUX-REIX, D. DELOMIER et M.-A. MOREL (eds.), *Frontières, du linguistique au sémiotique*. Limoges: Lambert-Lucas: 17-37.
- MOREL, M.-A. & DANON-BOILEAU, L. 1998. *Grammaire de l'intonation. L'exemple du français oral*. Paris – Gap: Ophrys.
- PORTES, C. 2004. *Prosodie et économie du discours: spécificité phonétique, écologie discursive et portée pragmatique de l'intonation d'implication*. PhD thesis. Université de Provence.
- RAYMOND, G. 2003. Grammar and Social Organization: Yes/No Interrogatives and the Structure of Responding. *American Sociological Review* 68 (6): 939-967.
- RELIEU, M. 1999. La réalisation et la réception du produit télévisuel comme accomplissements. In J.P. DESGOUTTE (ed.), *La mise en scène du discours audiovisuel*. Paris: L'Harmattan: 35-65.
- SACKS, H. 1992. *Lectures on Conversation*. Oxford – Cambridge: Blackwell.
- SACKS, H., SCHEGLOFF, E.A. & JEFFERSON, G. 1974. A Simplest Systematics for the Organisation of Turn Taking for Conversation. *Language* 50 (4): 696-735.
- SCHEGLOFF, E.A. 1968. Sequencing in Conversational Openings. *American Anthropologist* 70 (6): 1075-1095.
- SCHEGLOFF, E.A. 1991. Reflections on Talk and Social Structure. In D. BODEN & D.H. ZIMMERMAN (eds.), *Talk and Social Structure: Studies in Ethnomethodology and Conversation Analysis*. Cambridge: Polity Press: 44-71.
- SCHEGLOFF, E.A. 1996. Turn Organization: One Intersection of Grammar and Interaction. In E. OCHS, E.A. SCHEGLOFF & S.A. THOMPSON (eds.), *Interaction and Grammar*. Cambridge: Cambridge University Press: 52-133.

- SCHEGLOFF, E.A. 1998. Body Torque. *Social Research* 65 (3): 535-596.
- SCHEGLOFF, E.A. 2007. *Sequence Organization in Interaction: A Primer in Conversation Analysis*. Cambridge: Cambridge University Press.
- SELTING, M. 2000. The Construction of Units in Conversational Talk. *Language in Society* 29 (4): 477-517.
- TEN HAVE, P. 2007. *Doing Conversation Analysis: A Practical Guide*. London – Thousand Oaks – New Delhi: Sage Publications.
- WICHMANN, A. 2000. *Intonation in Text and Discourse. Beginnings, Middles and Ends*. Harlow: Longman.