The Intonation of Backchannel Tokens in Italian Collaborative Dialogues

Michelina Savino(✉)

Department of Education, Psychology, Communication,
University of Bari, Bari, Italy
michelina.savino@uniba.it

Abstract. This paper offers a contribution to the intonational modelling of backchannel lexical and non-lexical tokens in Italian, which can be used for improving naturalness in voice-based dialogue systems. Results of pragmatic and intonation analysis of five Map Task dialogues show that backchannel tokens can convey the intention of giving vs taking the floor by means of a rising vs falling terminal contour. However, they also indicate that this general rule can be override when other pragmatic and/or paralinguistic meanings need to be additionally conveyed.

Keywords: Backchannels · Acknowledgement tokens · Intonation · Map Task dialogues · Conversational moves · Turn-taking

1 Introduction

One of the still persisting challenges in man-machine interaction is the possibility of developing voice-based dialogue systems which can really help reproducing the naturalness of interaction between human beings. An important aspect of human speech communication whose modelling and implementation could strongly contribute to improve naturalness is represented by the use of backchannel or acknowledgement tokens (for example, Cathcart et al. 2003; Gravano and Hirschberg 2009; Misu et al. 2011). They include lexical and non-lexical tokens (like 'yes', 'yeah', 'mm', 'uh uh', and so on) which can be generally used for signalling that the listener is attending to the speaker and prompting her/him to go on, even though the range of their possible paralinguistic meanings and pragmatic functions can be wider (understanding, agreement, appreciation, assessment, passive recipiency, incipient speakership, etc. as mainly described in Conversation Analysis works, see for example Schegloff 1982; Jefferson 1983), also depending on communicative contexts. Since most of backchannels are non-lexical tokens, a crucial role in signalling those functions is played by intonation on its own, where intonational features can be obviously language-dependent. In this respect, a number of studies have been devoted to the description of the main prosodic characteristics of acknowledgement tokens in a number of languages (see for example Caspers 2000 for Dutch, Jurafsky et al. 1998, Ward 2004, Benus et al. 2007 for American English, Ward 1998 for Japanese), for linguistic-theoretical description aims as well as for modelling purposes in man-machine interaction. As far as Italian is concerned, one previous study has been

Z. Vetulani and J. Mariani (Eds.): LTC 2011, LNAI 8387, pp. 28–39, 2014.
DOI: 10.1007/978-3-319-08958-4_3

attested, which consists of a preliminary investigation on duration and tonal features of a wide range of short expressions including backchannels (Cerrato and D'Imperio 2003). In all these studies, analysis was carried out on task-oriented dialogues.

Aim of this paper is to provide a further contribution for determining the role of intonation in Italian backchannels, whose modelling can be useful for improving naturalness in Italian voice-based dialogue systems. Analysis is based on task-oriented dialogues, making results comparable with those obtained for other languages. Also, in eliciting data some parameters for controlling and enhancing backchanneling have been used, and a more "context-based" approach in interpreting backchannel functions has been adopted. Moreover, in classifying backchannels tokens with respect to the turn-taking dynamics, a "non-intonation based" operational definition is proposed, which avoids circularity in the pragmatic interpretation of such tokens as implying taking vs yielding the turn during interaction.

2 Materials and Methodology

2.1 Corpus

Spoken materials analysed consist of five Bari Italian dialogues elicited with a modified version of the Map Task method (Anderson et al. 1991), each having an average duration of 10–20 min, corresponding to the spoken productions of ten speakers. In a Map Task session, pairs of participants – an Instruction Giver (henceforth IG), and an Instruction Follower (henceforth IF) – is given a map. One of the two maps has a route drawn on it, and the task consists in reproducing as accurately as possible the route on the other map by exchanging information via the verbal channel. The task is complicated by the fact that the two maps are not identical in terms of presence and position of the landmarks, thus stimulating possible misunderstanding like in natural, everyday interaction. Differently from the original Map Task methodology, in Bari Italian sessions participants were not informed in advance that the two maps were different; neither they were told that the maps were identical, even though this is was they assumed (Grice and Savino 2003). In other words, before starting the task participants assumed they initially shared the same background knowledge (how this aspect has an influence on backchannel intonation will be discussed in Sect. 3.2).

In each recording session, eye contact was always inhibited, in order to maximise the use of the verbal channel during communication, including the production of verbal backchannels. In Map Task sessions, lack of eye contact stimulates the production of verbal backchannels in order to provide dialogue partners with increased verbal feedback (Boyle et al. 1994), presumably as a compensation for the missed non-verbal feedback normally conveyed by body gestures and gaze. Because of the impossibility of seeing each other, participants need to use more verbal feedback also for regulating turn-taking as much efficiently as possible. Besides the possibility of controlling these kinds of parameters during interaction, collaborative dialogues like Map Tasks are particularly suitable for studying and modelling intonation of backchannels for human-machine interaction (especially for voice-based dialogue

systems). In fact, in such interactional contexts the successfulness of information transferring is crucial for accomplish the assigned task, where the role of verbal feedback and turn-taking regulation is fundamental in making such an information exchange effective.

2.2 Pragmatic and Intonation Analysis

As a first step, a pragmatic analysis of the dialogues based on both orthographic transcripts and audio files listening was carried out. Such analysis consists of pragmatic annotation of utterances in terms of conversational moves, according to the Map Task coding scheme (Carletta et al. 1997). This scheme provides a broad category for describing backchannel phenomena, namely the ACKNOWLEDGE conversational move, defined as "[…] a verbal response that minimally shows that the speaker has heard the move to which it responds, and often demonstrates that the move was understood and accepted" (Carletta et al. 1997, p. 19).

Since we were interested in determining possibly specific intonational cues used in backchannels for regulating turn-taking during interaction, in the pragmatic analysis we included a further distinction (introduced by Jefferson 1983 and used later by Jurafsky et al. 1998) between:

- ACKNOWLEDGE tokens reflecting Passive Recipiency (henceforth PR), also called continuers, acknowledging that the other speaker still has the turn;
- ACKNOWLEDGE tokens reflecting Incipient Speakership (henceforth IS), indicating the intention of taking the floor, reflecting "[…] preparedness to shift from recipiency to speakership" (Jefferson 1983, p. 4).

Decision on whether an ACKNOWLEDGE token could be classified as an example of PR or IS was based on whether a change of speaker occurred after that token or not. Following Cathcart et al. (2003), we identified Transition Relevance Places (henceforth TRPs, defined as points for potential turn switching between conversational partners, Sacks et al. 1974) at move boundaries. When a change of speaker was observed after the ACKNOWLEDGE token/move, that token was categorised as an ACKNOWLEDGE_PR, like in the following excerpt:

IG: vai verso destra
('go rightwards')
INSTRUCT
IF: sì
('yes')
ACKNOWLEDGE_PR
IG: poi risali
('then go up again')
INSTRUCT

When after an ACKNOWLEDGE token/move the same speaker went on speaking (i.e., no change of speaker was involved after backchanneling), that token was labelled

as an ACKNOWLEDGE_IS, like in the following excerpt (bars indicate move boundaries within the same turn):

IG: devi passare sempre alla sinistra del bar
('you have to go through the left side of the bar')
INSTRUCT
IF: sì || ma il bar io me lo lascio sulla sinistra?
('yes' || ' but the bar, shall I leave it from the left side?')
ACKNOWLEDGE_IS || QUERY_YN
IG: sì sì sì
('yes yes yes')
REPLY_YN

This operational criterion was adopted in order to avoid circularity in the pragmatic interpretation of backchannels in turn-taking, i.e. using intonation contour for classifying these tokens as reflecting PR or IS when the scope of the study is *deriving* their intonational characterisation with respect to PR and IS.

Items selected for intonation analysis are all lexical and non-lexical (monosyllabic and bisyllabic) tokens used in ACKNOWLEDGE moves, such as 'sì' (yes), 'mm', 'okay' (this English token is normally used as such by Italian speakers), 'eh', 'aha', etc., for a total amount of 463 tokens. All tokens have been intonationally analysed in terms of overall F0 shapes (fall, rise, fall-rise, etc.), basing on both perceptual judgement and F0 inspection. Pragmatic and intonation annotations have been carried out by two independent labellers (inter-labellers' agreement score > 89 %) using the Praat software tool for speech analysis (Boersma and Weenink 2001). Results discussed in this paper refer to the mostly occurring acknowledgement tokens in our Bari Italian dialogues, namely 'sì', 'mm', 'okay', 'eh' (342 tokens).

Note that these tokens are also found in the dialogues as positive replies (REPLY_Y conversational moves) to yes-no questions (QUERY_YN, CHECK and ALIGN conversational moves). An intonational characterisation of these tokens in relation to the two different pragmatic functions (acknowledgements vs positive replies) is described in Savino (2010), and preliminary results on prosodic features useful for their pragmatic disambiguation are provided in Savino and Refice (2013).

3 Results and Discussion

3.1 Backchannels and Turn-Taking

In modelling backchannels and their relation to turn-taking, it could also be useful to determine whether token type choice might be a parameter involved in such a dynamics, as suggested by Conversational Analysis (for example, Jefferson 1983). Figure 1 shows the distribution (in percentage) of the four main token types ('sì', 'mm', 'okay', 'eh') with respect to PR vs IS conveyed in backchanneling. To the extent of the statistical significance of our data, results show that in our dialogues Italian tokens 'sì' and 'mm' are mostly used for reflecting PR in backchanneling (i.e.

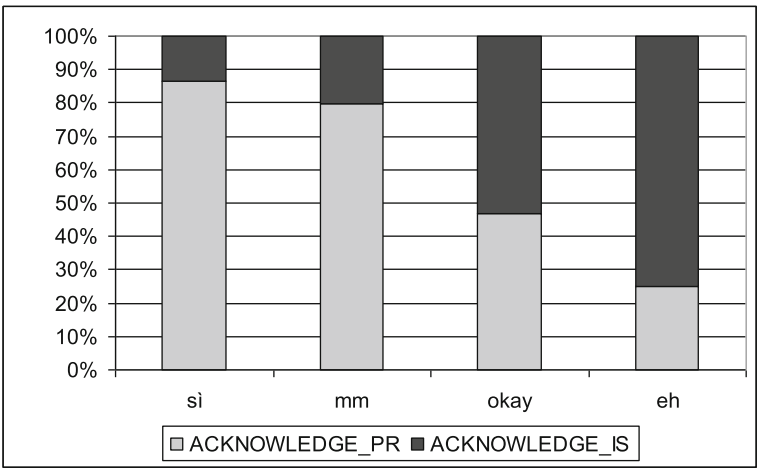

Fig. 1. Distribution of the four main token types across the two ACKNOWLEDGE modalities, i.e. ACKNOWLEDGE_PR (Passive Recipiency; current speaker does not take the turn after acknowledging) and ACKNOWLEDGE_IS (Incipient Speakership, current speaker takes the floor after acknowledging).

when speakers do not take the turn afterwards), token 'eh' is mostly correlated with IS (i.e. when speakers do take the floor after backchanneling), whereas 'okay' seems to be almost equally used in both cases. These results suggest for Italian a possibly different preference in ACKNOWLEDGE token choice in relationship to PR vs IS in comparison to English. According to Jefferson (1983), in fact, in English 'mm' is mostly used with PR, whereas 'yes'/'yeah' is mainly associated with IS. Even though more statistical data is needed to confirm this preliminary outcome, it nevertheless indicates that token choice is an important language-specific and cultural parameter to take into account when modelling backchannels, especially for human-machine interaction.

As to the intonation analysis, Fig. 2 shows the overall distribution (in percentage) of intonation contours with rising vs falling terminal contour across the two main token types, i.e. those implying PR (ACKNOWLEDGE_PR) and those reflecting IS (ACKNOWLEDGE_IS). It can be observed that tokens involving taking the floor after backchanneling are predominantly characterised by a falling terminal contour, whereas the F0 shape of tokens implying not taking the turn after backchanneling mostly ends with a rise. These results suggest a specific role of intonation in signalling turn-taking during backchanneling.

3.2 Passive Recipiency and Pragmatic Context

Results presented in Sect. 3.1 above are also compatible with the general meaning of "openness, non-conclusiveness, continuity" attributed to rising/high pitch, and the opposite meaning to the falling/low pitch, as enunciated by the frequency code theory

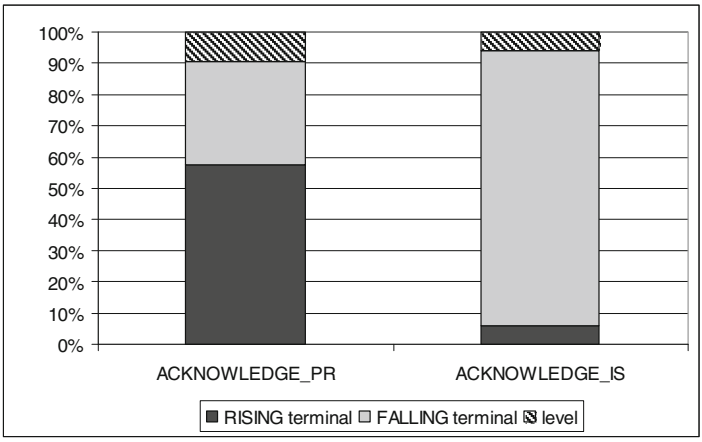

Fig. 2. Distribution of rising vs falling terminal contours across the two ACKNOWLEDGE modalities, i.e. Passive Recipiency (PR) and Incipient Speakership (IS).

(Ohala 1984). On the other hand, rising terminal backchannels functioning as continuers (i.e. involving PR) have been attested also in other languages (see refs. cited in Sect. 1). However, our statistics show that, differently from the clear-cut intonational characterisation of acknowledgement tokens implying IS, that of tokens involving PR shows some variability, since around 30 % of cases have a falling instead of a rising terminal F0 pattern. It can be hypothesised that in Italian backchannel continuers, the intonational choice could be influenced by additional pragmatic meanings "triggered" by specific pragmatic contexts. In order to verify this hypothesis, we identified these interactional pragmatic contexts as the conversational move pairs consisting of the ACKNOWLEDGE_PR token/move produced by the current speaker and the move realised by dialogue partner in his/her immediately preceding turn (when the partner's turn consisted of more than one conversational moves, only the last one was considered for the analysis). Figure 3 shows the distribution of falling vs rising terminals in ACKNOWLEDGE_PR tokens as a function of different interactional contexts, namely whether the dialogue partner's immediately preceding turn realised an INSTRUCT (giving instructions), an EXPLAIN/CLARIFY (giving explanations or clarifications), a REPLY_W (replying to a wh-question), a REPLY_YN (giving a positive or negative answer to a yes-no question), or an ACKNOWLEDGE move. It can be noted that the PR tokens with a terminal rise are predominantly concentrated (>80 % of cases) as following INSTRUCT moves, whereas those characterised by a falling terminal contour *always* occur as feedbacks to ACKNOWLEDGE moves, and as backchannel tokens immediately produced after EXPLAIN/CLARIFY, REPLY_W, and REPLY_YN moves in the majority of cases.

We also looked at the possible correlation between the terminal contour type of ACKNOWLEDGE_PR tokens and that of the above mentioned dialogue partners' immediately preceding moves. Such a distribution is given in Table 1, where it can be observed that PR backchannel tokens ending with a rising F0 contour are

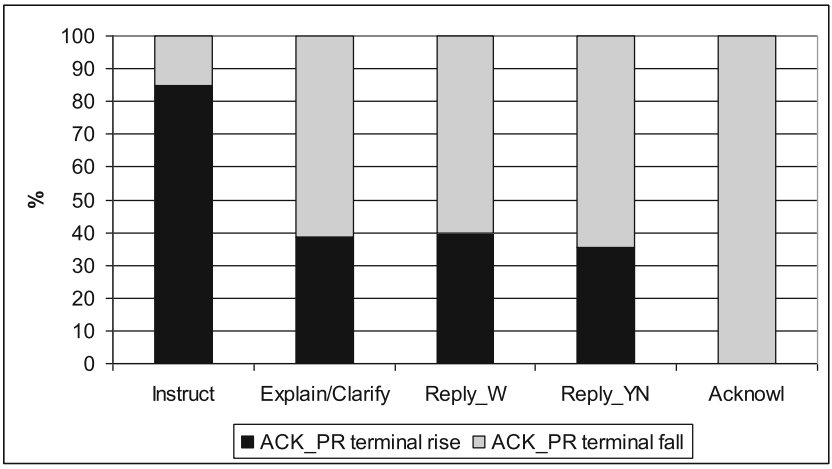

Fig. 3. Distribution of rising vs falling terminal contours in ACKNOWLEDGE_PR tokens as a function of the dialogue partner's conversational move realised in his/her turn immediately preceding the ACKNOWLEDGE_PR token.

Table 1. Distribution of rising vs falling terminal contours in ACKNOWLEDGE_PR tokens as a function of the dialogue partner's conversational move and relating terminal contour type realised in his/her turn immediately preceding the ACKNOWLEDGE_PR token.

Preceded by→	INSTRUCT terminal rise (%)	INSTRUCT terminal fall (%)	EXPLAIN/REPLY/ ACKNOWLEDGE terminal rise (%)	EXPLAIN/REPLY/ ACKNOWLEDGE terminal fall (%)
ACKNOWLEDGE_PR terminal rise	90	0	1	9
ACKNOWLEDGE_PR terminal fall	0	37	0	63

predominantly preceded by partners' conversational moves marked by a rising terminal as well. On the other hand, PR tokens with terminal falls function mostly as feedback to dialogue partners' immediately preceding moves being also characterised by a terminal fall. These results indicate that the typical use of a rising terminal contour in backchannels with PR can be conditioned by additional pragmatic and communicative meanings to be conveyed beyond that of turn-taking disposition.

3.3 Backchannels, Common Ground and Speaker Attitude

Outcomes discussed so far on ACKNOWLEDGE_PR tokens have been focussed on the distribution of contour types as grouped into two overall categories: F0 contours ending with either a rise or a fall. In order to shed more light on the range of possible additional meanings which can be conveyed by intonation in these backchannels, an overview of specific intonation patterns encountered in the dialogues are presented

and discussed. As shown in Fig. 4, for the F0 terminal rising backchannels a wide range of contours is available: rise, fall-rise, stylised rise, and high rise (note that in the figure, rise and high rise are collapsed in one category). In the dialogues, it has been observed that rise, fall-rise and stylised-rise contours are normally found in tokens signalling understanding/agreement about information (typically, instructions) being received by the dialogue partner, i.e. in backchannels following INSTRUCT moves as described in Sect. 3.2.

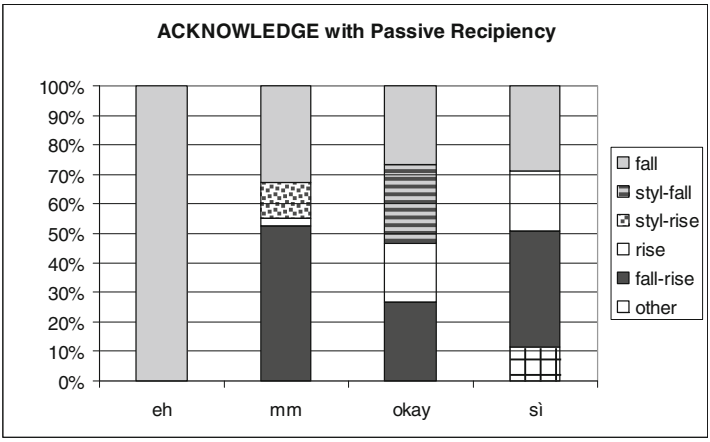

Fig. 4. Distribution of intonation contour types across the four token types as realised in ACKNOWLEDGE_PR moves.

The rising contour seems to be the "default" F0 shape for ACKNOWLEDGE_PR tokens. An example of token 'sì' with a rising contour is shown in Fig. 6(a). It has also been observed that the fall-rise and the stylised rise are used by dialogue participants when transferred information – typically via instructions about presence and/or position of landmarks – are fully compatible with features on her/his own map. In other words, they are used for confirming the current status of common background knowledge (common ground, see for example Clark and Schaefer 1989; Clark and Brennan 1991), and for conveying a consequently positive speaker's attitude. In particular, the stylised rise is only found in 'mm' tokens, and in this case they are produced as bisyllabic, as also observed in English backchannels (Ward 2004). The stylised rise is realised as a pitch step-up from the first to the second syllable, where the pitch excursion is not wide. An example of a 'mm' token with stylised rise is shown in Fig. 5(b), whereas Fig. 5(a) offers an example of 'mm' with fall-rise (in this case the token is realised as bimoraic).

A number of backchannels with a high rising pitch contour were also encountered, typically produced by IFs at/around the end of a set of instructions for completing a (sub)task. This contour type has been already described in Bari Italian for signalling pre-finality, i.e. marking the antepenultimate item in a sequence (Savino 2001, 2004).

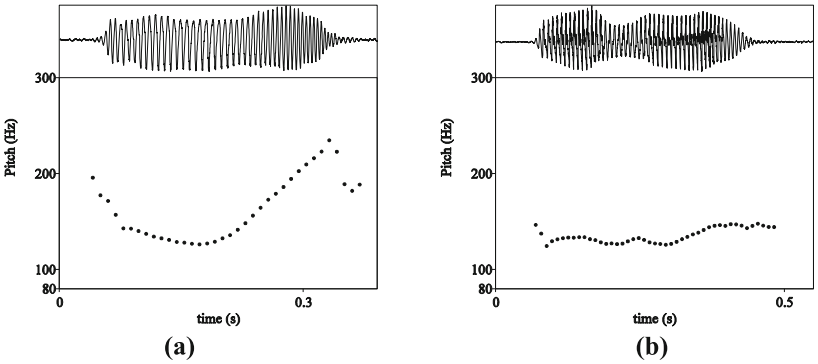

Fig. 5. ACKNOWLEDGE_PR 'mm' realised as a fall-rise (a), and as a stylised rise (b). Note that in the first case the token is realised as bimoraic, in the second as bisyllabic.

This suggests that in ACKNOWLEDGE_PR tokens, this contour conveys speaker's understanding/agreement that the end of a (sub)set of instructions for completing a (sub)task is approaching. An example of such a contour is shown in Fig. 6(b).

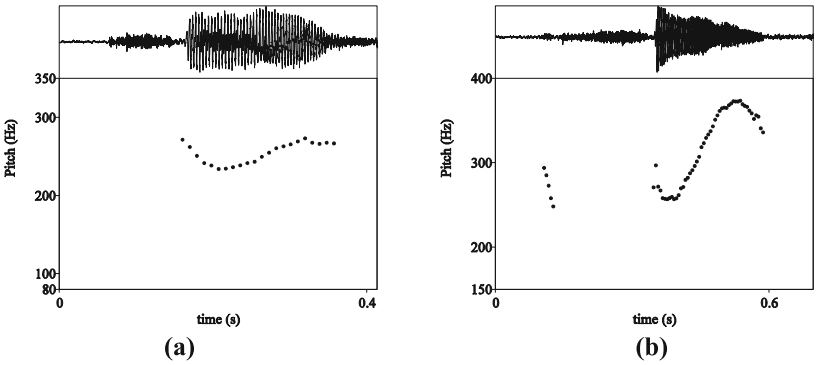

Fig. 6. ACKNOWLEDGE_PR 'sì' realised with a rise (a), and with a high rise (b). Note that in the latter case it signals that the end of a (sub)task is approaching

As to the ACKNOWLEDGE_PR tokens realised with a falling terminal, intonation analysis reports cases of fall and stylised fall melodic shapes (Fig. 4). By looking at the dialogue contexts where a falling rather than the typical rising terminal contour is encountered, we observed that fall and stylised fall F0 patterns are usually found where participants have not discovered yet that the two maps are different. Such contours are typically produced in backchannels by IFs when receiving instructions for drawing the route which are not compatible with presence and/or position of landmarks on her/his own map. Even though the use of backchannels generally implies signalling understanding and agreement, in these cases a falling contour seems to convey disagreement or disappointment about what is assumed to be the currently

shared background knowledge. Therefore, a falling contour seems to signal a *negative* backchanneling, as there is a mismatch between participants' assumed common background knowledge at that moment of the interaction.

Some cases of fall and stylised fall in acknowledgement tokens with RP are also found at the last turn(s) of the dialogue. Because of their position within the dialogue, and the melodic shape typically associated the meaning of "finality, conclusiveness", these backchannels can be interpreted as signalling understanding/agreement that the task has been completed. Some other cases of such falling contours are typically found in tokens acknowledging interlocutor's preceding acknowledgment, as already shown in Sect. 3.2 above. This outcome is particularly interesting, as the occurrence of this pragmatic type of backchannels contrasts with the formal definition of these token types proposed by Ward and Tsukahara (2000). According to these authors, backchannels do not require acknowledgement by the other speaker, yet our results show that this is possible, demonstrating that backchannel strategies can largely vary, depending on individual, cultural and above all communicative context variables. Examples of a 'mm' token realised with a fall (a) and 'okay' with a stylised fall (b) are shown in Fig. 7. Note that the stylised falls are intonationally realised as an F0 step-down from the first to the second syllable, with a relatively reduced pitch excursion.

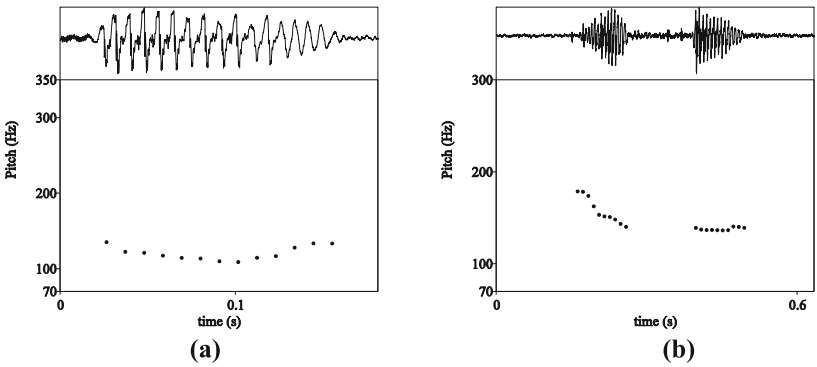

Fig. 7. ACKNOWLEDGE_PR 'mm' realised with a falling contour (a), and 'okay' with a stylised fall (b). The former is observed when the speaker finds out a mismatch in assumed shared background knowledge, the latter at the end of the whole task.

Finally, the distribution of intonation contours across the four main ACKNOWLEDGE_PR tokens (Fig. 4) shows that the token 'eh' is always realised with a falling F0 pattern. In this case, there is a specific choice not only in terms of melodic shape, but also in terms of token type for cueing information status, speaker's attitude and discourse structure while acknowledging understanding, as described above for falling terminal backchannels.

The intonational choices of ACKNOWLEDGE_PR tokens in relation to the pragmatic functions and attitudinal meanings discussed above are schematised in Table 2.

Table 2. Intonation contours and their association with pragmatic functions/paralinguistic meanings in ACKNOWLEDGE_PR tokens as observed in Bari Italian Map Task dialogues (BK = Background Knowledge).

Pragmatics	rise	high rise	fall-rise	styl-rise	fall	styl-fall
Acknowledges understanding/agreement on received info ("default")	**x**					
Confirms current status of assumed common BK			**x**	**x**		
Signals end of instructions set for completing a (sub)task is approaching		**x**				
Signals disagreement or disappointment on assumed currently shared BK					**x**	
Acknowledges dialogue partner's previous acknowledgement					**x**	**x**
Signals end of (set of instructions for completing) the (sub)task					**x**	**x**

4 Conclusions

Results from pragmatic and intonation analysis of Bari Italian Map Task dialogues have provided a number of intonational features for modelling backchannel or acknowledgement tokens which can be useful for improving naturalness in speech dialogue systems for Italian. They have shown that when speakers do not take the floor after backchanneling, they produced acknowledgement tokens predominantly characterised by a rising terminal intonation pattern. On the other hand, when speakers take the turn right after backchanneling, their acknowledgement tokens have a falling terminal instead. Therefore, in Italian the intention of taking the floor or not while backchanneling seems to be conveyed intonationally by a falling vs rising terminal F0 contour. However, analysis also suggests that the general "rule" of a terminal rising for signalling the intention of yielding the floor might not be applied when additional co-occurring pragmatic and/or paralinguistic meanings need to be conveyed, like for example the status of participants' assumed shared background information at that time of interaction, or the current stage of interaction with respect to the end, i.e. the accomplishment of the task.

References

Anderson, A., Bader, M., Bard, E.G., Boyle, E., Doherty, G., Garrod, S., Isard, S., Kowtko, J., MacAllister, J., Miller, J., Sotillo, C., Thompson, H., Weinert, R.: The HCRC map task corpus. Lang. Speech **34**(4), 351–366 (1991)

Benus, S., Gravano, A., Hirschberg, J.: The prosody of backchannels in American English. In: 16th International Congress of Phonetic Sciences, Saarbruecken, pp. 1065–1068 (2007)
Boersma, P., Weenink, D.: Praat, a system for doing phonetics by computer. Glot Int. **5**(9/10), 341–345 (2001)
Boyle, E., Anderson, A.H., Newlands, A.: The effects of visibility on dialogue and performance in a cooperative problem solving task. Lang. Speech **37**(1), 1–20 (1994)
Carletta, J., Isard, A., Isard, S., Kowtko, J., Doherty-Sneddon, G., Anderson, A.: The reliability of a dialogue structure coding scheme. Comput. Linguist. **23**(1), 13–32 (1997)
Caspers, J.: Melodic characteristics of backchannels in Dutch Map Task dialogues. In: International Conference of Spoken Language Processing 2000, Bejing (2000)
Cathcart, N., Carletta, J., Klein, E.: A shallow model of backchannel continuers in spoken dialogue. In: 10th Conference of the European Chapter of the ACL, vol. 1, pp. 51–58 (2003)
Cerrato, L, D'Imperio, M.: Duration and tonal characteristics of short expressions in Italian. In: 15th International Congress of Phonetic Sciences, pp. 1213–1216 (2003)
Clark, H.H., Brennan, S.E.: Grounding in communication. In: Resnick, L.B., Levine, J., Tearsely, S.D. (eds.) Perspectives on Socially Shared Cognition, pp. 127–149. American Psychological Association, Washington (1991)
Clark, H.H., Shaefer, E.F.: Contributing to discourse. Cog. Sci. **13**, 259–294 (1989)
Gravano, A., Hirschberg, J.: Backchannel-inviting cues in task-oriented dialogue. In: Interspeech 2009, pp. 1019–1022 (2009)
Grice, M., Savino, M.: Map tasks in Italian: asking questions about given, accessible and new information. Catalan J. Linguist. **2**, 153–180 (2003)
Jefferson, G.: Notes on systematic deployment of the acknowledgment tokens "yeah" and "mm hm". TILL, Tilburg University, The Netherlands (1983)
Jurafsky, D., Shriberg E., Fox, B., Curl, T.: Lexical, prosodic, and syntactic cues for dialogue acts. In: ACL/COLING Workshop on Discourse Relations and Discourse Markers, pp. 114–120 (1998)
Misu, T., Mizukami, E., Shiga, Y., Kawamoto, S., Kawai, H., Nakamura, S. (2011). Towards contruction of spoken dialogue system that evokes users' spontaneous backchannels. In: SIGDIAL 2011, Portland, pp. 259–265
Ohala, J.: An ethological perspective on common cross-language utilization of F0 of voice. Phonetica **41**, 1–16 (1984)
Sacks, H., Schegloff, E.A., Jefferson, G.: A simplest systematics for the organization of turn-taking for conversation. Language **50**(4), 696–735 (1974)
Savino, M.: Non-finality and pre-finality in Bari Italian intonation: a preliminary account. In: Eurospeech 2001, Aalborg, pp. 939–942 (2001)
Savino, M.: Intonational cues to discourse structure in a variety of Italian. In: Gilles, P., Peters, J. (eds.) Regional Variation in Intonation, pp. 145–159. Niemeyer, Tübingen (2004)
Savino, M.: Prosodic strategies for backchanneling in Italian Map Task dialogues. In: ISCA Tutorial and Research Workshop on Experimental Linguistics, Athens, pp. 157–160 (2010)
Savino, M., Refice, M.: Acknowledgement or reply? Prosodic features for disambiguating pragmatic functions of the Italian token 'sì'. In: SpeD 2013, Cluj-Napoca, pp.117–122 (2013)
Schegloff, E.: Discourse as an interactional achievement: some uses of 'hu huh' and other things that come between sentences. In: Tannen, D. (ed.) Analyzing Discourse: Text and Talk, pp. 71–93. Georgetown University Press, Washington DC (1982)
Ward, N.: The relationship between sound and meaning in Japanese backchannels grunts. In: Meeting of the Japanese Association for National Language Processing, Fukuoka, pp. 464–467 (1998)
Ward, N.: Pragmatic function of prosodic features in non-lexical utterances. In: Speech Prosody 2004, Nara, (on CD-ROM) (2004)
Ward, N., Tsukahara, W.: Prosodic features which cue back-channel responses in English and Japanese. J. Pragmat. **32**, 1177–1207 (2000)