

# Addressee backchannels steer narrative development

Jackson Tolins<sup>\*</sup>, Jean E. Fox Tree<sup>\*</sup>

Psychology Department, Social Sciences II, University of California, Santa Cruz, CA 95064, United States

Received 29 January 2014; received in revised form 10 June 2014; accepted 16 June 2014



## Abstract

Brief addressee responses such as *uh huh*, *oh*, and *wow*, which are called *backchannels*, are typically considered reactive phenomena – devices that respond in various ways to what was just said. Addressees, in providing backchannels, actively shape story telling in spontaneous dialogue (Bavelas et al., 2000). We contrasted generic backchannels with context-sensitive specific backchannels within a collection of face-to-face dialogues and in a narrative completion experiment. The analysis demonstrates that storytellers respond in distinct patterns to the two categories of backchannels. After generic backchannels, they provide discourse-new events. After specific backchannels, they provide elaborative information on previously presented events. Results from an experiment support this analysis, indicating that people reading transcripts of the conversation predict a similar pattern of story continuation following generic versus specific backchannels. We conclude that addressee responses are not only reactive, but proactive and collaborative in the shaping of narrative.

© 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

**Keywords:** Backchannels; Dialogue; Narrative; Addressee; Collaborative language

## 1. Introduction

When people tell stories to one another, as is common in spontaneous conversation, one conversational partner frequently speaks for extended periods, during which the other interactant can, and often does, provide a variety of comments on the story. These *backchannels* include verbal responses, such as *yeah*, *oh*, *okay*, or *mhm*, and visual displays, such as facial expressions, nods, and gestures (Bavelas and Gerwing, 2011; Bertrand et al., 2007; Yngve, 1970). Transcript (1) presents an example story telling, in which a student, S2, described a cinema course that he was enrolled in. As S2 described a movie-watching event to his addressee, S1, S1 actively participated in the interaction, providing three instances of verbal backchannels, in lines 7, 10, and 13 (all transcripts are of speech collected in our laboratory and are presented in broad Jeffersonian transcription).

(1)

- 1 S2: We watched a movie called Chun King Express last night  
2 S1: Oh ya. I've=  
3 S2: =It was crazyness  
4 S1: Did you like it?  
5 S2: It was ki:nd of intense they- they set it up i:n like- how it's like meant

<sup>\*</sup> Corresponding authors. UC, Santa Cruz, 1156 High St., Santa Cruz, CA 95064, United States. Tel.: +1 612 802 9067; fax: +1 831 459 3519. E-mail addresses: [jtolins@ucsc.edu](mailto:jtolins@ucsc.edu) (J. Tolins), [foxtree@ucsc.edu](mailto:foxtree@ucsc.edu) (J.E. Fox Tree).

- 6 to be watched so it was like <36 millimeter> or something like that  
 7 S1: Mhm  
 8 S2: And if you do it like that you gotta do like all the different reels and  
 9 you gotta connect. u:m  
 10 S1: Mhm  
 11 S2: I think they like left out a reel or something cause the movie like  
 12 completely didn't make sense at all [and was all like-  
 13 S1: [Really

From a unilateral perspective on language processing, in which comprehension and production are seen as distinct and isolated processes, backchannels are likely to be viewed as unnecessary, or at best superfluous. Indeed, a number of studies of backchannels have used optionality as a key definitional criterion (e.g. Ward and Tsukahara, 2000). Typically, research in this vein has focused on backchannels as a means of signaling turn taking goals – specifically, as a means to avoid taking over the floor from the current speaker. This has led to a view of backchannels as supportive, but not central. They are, in essence, a secondary message, as the label *backchannel* implies. In this conceptualization, backchannels have also been referred to as *reactive tokens* (Clancy et al., 1996), *response tokens* (Gardner, 2001), and *accompaniment signals* (Kendon, 1967). Addressees are seen as passive recipients of information, with backchannels being used to display addressees' acceptance of speakers' planned multi-turn utterances. We will refer to theories of backchannels within these paradigms as *reactive backchannelling theory*.

Another conceptualization of backchannels is that they are central to conversational success, demonstrating the producer's active participation in not just turn taking, but in the development of the speaker's talk. In dialogue focused on joint activities such as referential card tasks or building models, backchannels serve as *project markers* of particular types: *acknowledgement tokens*, *agreement tokens*, or *consent tokens* – each of which makes different comments on the ongoing talk (Bangerter and Clark, 2003). Acknowledgement tokens such as *uh huh* recognize what the speaker said as a contribution to the conversation, agreement tokens such as *right* indicate alignment with the speaker's position, and consent tokens such as *okay* indicate agreement to a joint plan of action. By providing a particular token at a particular point in the interaction, the addressee actively steers the ongoing collaborative task in a particular direction. A speaker's role involves not only talking, but actively monitoring addressee's backchannel communications as a means for altering his or her own talk in a precisely timed manner (Clark and Krych, 2004). In this conceptualization, addressees are active participants in the joint construction of spontaneously developing dialogue, which we will call the *proactive backchannelling theory*.

For the proactive backchannelling theory, addressee behaviors are actively involved in the unfolding activity. At the same time, speakers actively monitor addressees for these responses and adjust their talk accordingly (Clark and Murphy, 1982; Clark and Krych, 2004). This holds true not only for explicitly task-oriented dialogues but narration as well. Storytellers may take up their addressees' backchannels in a number of ways, ratifying and incorporating these responses into the development of the narrative (Norrick, 2010a,b, 2012). When addressee responses are controlled experimentally, the types of backchannels provided to the speaker shape the narrative content (Bavelas et al., 2000). In dyads where addressees did not provide context-specific assessments such as *wow* or nonverbal displays such as grimacing, speakers told qualitatively worse stories with significantly less climactic endings. In a similar study, addressee affective displays, such as smiling or frowning in response to the speaker, modulated the level of abstract language present in the speaker's talk (Beukeboom, 2009). So in both explicitly goal-directed, object-oriented tasks and in narrative story telling, backchannels function beyond simply responding to previous talk or signaling acceptance of a planned multi-utterance speaker turn. Instead, addressee behaviors are involved in the moment-by-moment collaborative production of talk.

The present report extends the study of the proactive role of backchannels in co-telling, focusing on spontaneous narratives occurring in the context of face-to-face conversation. From an inductive and qualitative analysis of conversation, we establish specific hypotheses about the relation between backchannel types and speakers' continuing talk, which we then test using an experimental paradigm. While the influence of backchannels on speaker talk has been previously explored at a more global level of narrative analysis (Bavelas et al., 2000; Beukeboom, 2009), the current in-depth analysis of conversational sequences coupled with experimental findings show how backchannels affect the discourse-level development of the directly subsequent talk.

## 2. Perspectives on backchannels

The study of backchannel communication strategies has a long history (see e.g. Dittmann and Llewellyn, 1986; Duncan and Fiske, 1977; Fries, 1952; Yngve, 1970). Across this literature, continued research has been motivated by an interest in what types of information backchannels provide. Research on backchannel communication has focused primarily on two aspects, functional distinctions between different categories and the organized placement of backchannels within the sequential conversational structure.

## 2.1. Function

The first vein of research spans a variety of paradigms investigating differences in what backchannels display. One categorical distinction is between *specific* and *generic* backchannels (Goodwin, 1986; Bavelas et al., 2000), also called *assessments* and *continuers* respectively (Goodwin, 1986; Stivers, 2008). Specific backchannels, such as *oh wow*, are context sensitive in that they express addressees' responses to the content of the previous turn. Generic backchannels, such as *uh huh* or *yeah*, respond not to the content of the previous talk, but rather to the need to display understanding and continued attention to the speaker.

Of course, it is possible to produce generic backchannels so that they imply commentary on the preceding utterance: imagine producing *uh huh* with an elongated vowel or rise in pitch (cf. Tomlinson and Fox Tree, 2011) or saying *yeah* with a tone of uncertainty or enacted surprise. We propose that in these cases, the added intonational information changes what is generally considered a generic backchannel to take on the meaning of a specific backchannel. Without the added prosodic cues (or possibly other kinds of cues, such as visual cues) the generic backchannels would serve only as grounding displays rather than commentary.

Bangerter and Clark (2003) presented a similar functional distinction to the generic/specific, continuer/assessment distinctions. They focused on backchannels as *project markers* used to coordinate transitions across joint activities. Through the analysis of task-oriented collaborative dialogue in which the interactants accomplished a set of hierarchically nested tasks, they found that backchannels like *uh huh* did not demonstrate that the addressee did not wish to take a turn. Instead, they displayed the addressee's acknowledgement of a speaker's talk as proposing or continuing a particular joint action. This function is contrasted with other backchannels, such as *okay* and *alright*, which primarily marked the completion and transition out of a particular joint project or subproject. Importantly, the perspective here is not one of structuring the organization of the conversation itself, but rather analyzing the dialogue as a means through which joint activities are accomplished.

Together, these studies illustrate the variety of functional perspectives taken toward backchannel communication, and the categorical distinctions based on actions achieved, cutting across type and modality. Indeed, backchannels may accomplish many things at once, at different levels of analysis. Brunner (1979) suggests three in his analysis of smiles as backchannels: (1) backchannels signal involvement and participation in a joint activity, (2) backchannels signal understanding, or lack thereof, and (3) backchannels signal the addressee's affective or informational response to the speaker's talk and affiliation with the speaker's presented stance. He further demonstrated that smiles accomplish actions at all three levels. Similarly, Clark and Krych (2004) suggested that backchannels were used to display addressee uptake at four levels of joint action (Clark, 1996): (a) attending, (b) identifying, (c) understanding, and (d) compliance. Attending and identifying are subsumed in comprehension, consisting of displaying attention and word identification. Understanding and compliance represent higher levels built on the first, representing integration of meaning into the discourse and context, and the acceptance of the proposed conversational action of the speaker's talk. In line with much prior research, we take generic backchannels as signals of participation and understanding, and specific backchannels as signals of the addressee's stance toward the content of the speaker's talk.

## 2.2. Placement

The second vein of research centers on investigations of where within the speaker's talk backchannels occur. This perspective is driven by a focus on turn taking and the structural organization of interaction. Whatever the functional role of a particular backchannel, each is seen as important in the management of selecting who will speak next (Sacks et al., 1974). Following from this paradigm, research has also been conducted with the goal of providing an analysis of what cues within the speaker's talk – whether prosodic, syntactic, or embodied – act as invitations for the addressee to provide a backchannel (see e.g. Koiso et al., 1998; Morency et al., 2010; Bavelas et al., 2002).

In consideration of turn-taking and related phenomena, backchannels have been viewed as markers involved in indicating which participants within a conversation may hold the floor next (Duncan, 1972, 1974; Duncan and Fiske, 1977; Sacks et al., 1974). Because of their involvement in the sequential organization of a conversation, backchannels are said to occur typically at places wherein transition from one speaker to the next is particularly relevant, *transition relevance places* (Sacks et al., 1974). Here again differences between generic and specific backchannels are present. In his analysis comparing the sequential organization surrounding backchannels, Goodwin (1986) presented examples indicating that while continuers occur between two units of talk, assessments most commonly occur within a single unit, or turn, of speaker talk. As displays of continued attention, backchannels such as *uh huh* and *mhm* were viewed as an addressee's explicit agreement that the current speaker can engage in a multi-turn utterance (Schegloff, 1982).

An extension of this emphasis on turn coordination has been the analysis of cues within a speaker's talk that invite backchannels. Duncan and Fiske's (1977) signal-based theory of turn taking considered features in the speaker's talk that occurred immediately before speaker role changes as well as those preceding backchannels. Typical cues taken into

consideration as requests for backchannels are those that occur systematically at the ends of turns. For example, [Ward and Tsukahara \(2000\)](#) focused on a period of low prosody found at the end of utterances prior to addressee responses in both English and Japanese. Similarly, syntactic cues that are taken as invitations of backchannels are typically those that mark syntactic structures as complete such as the part of speech of final morphemes ([Koiso et al., 1998](#)). Because of the focus on finding particular locations within speaker talk in which backchannels occur, categorical distinctions across types of backchannels in relation to particular speaker cues have not been emphasized.

One potentially problematic aspect of the turn-management function is that what makes a backchannel a backchannel may only be determined in retrospect, once the conversation has concluded. Logically, while it is being produced, a backchannel could be the start of a turn, as illustrated transcript (2). Transcript (2) was collected during a tangram referential card task in which two participants worked together to negotiate descriptions for abstract shapes (corpus described in [Fox Tree, 1999](#)). As S3 described which of three cats S4 should select, S4 responded with backchannels in lines 17, 19, and 21.

- (2)
- 14 S3: ok I think I've got three that look like cats too u:m like one's a cat  
 15 that's kinda lying down and one a cat that's standing up and one's  
 16 a cat that's <kindof> three forty-five degree angle?  
 17 S4: uh huh  
 18 S3: 's tha' right?  
 19 S4: yup  
 20 S3: well this is the one that's kindof forty-five degree angle  
 21 S4: okay  
 22 S3: sort of sort of standing up and bending over?  
 23 S4: yeah

With this snippet of the conversation, S4's *yeah* in line 23 appears as if it may be a backchannel. That is, as the talk is unfolding, at the moment S4's *yeah* is spoken, S3 might reasonably hear it as a backchannel and continue talking. Had this happened, line 23 would be considered a backchannel. What actually happened, however, is displayed in (3).

- (3)
- 22 S3: sort of sort of standing up and bending over?  
 23 S4: yeah, like looking at something?  
 24 S3: yeah! exa- with eyes like it's looking at you  
 25 S4: nkay

A similar argument was made by other researchers who observed that turn-management functions frequently ascribed to some spontaneously produced phenomena may be epiphenomena of other functions. For example, Fox Tree and Schrock wrote, "*You know* or *I mean* may fall at the beginning, middle, or end of a turn for reasons unrelated to turn management" (2002:732). Others go further in arguing that turn-taking rules themselves have not been sufficiently substantiated. O'Connell, Kowal, and Kaltenbacher argued, "The retrospective assertion that *N* positions in a given conversation are TRPs [transition relevance places], (i.e., were somehow relevant for turn-taking), where, let us say, *N* – 34 have been used for taking turns, is a meaningless *post factum* intellectual exercise" (1990:351) and that "Kinesics, prosody, content, knowledge and attitude of the interlocutors about the topic, about one another, and about the situation – all these situational elements can change the direction and pace of turn-taking from moment to moment. Neglect of them renders the simplest systematics completely sterile" (1990:360).

In addition, most approaches to the placement of backchannels have focused on verbal backchannels, and particularly short single word tokens, which are much more likely to occur in places that do not overlap with the speaker's talk, thus emphasizing their role in turn coordination. When we expand our notion of backchannel communication to include non-verbal behaviors such as smiles, nods, or affective displays, we see that backchannels along these modalities do not adhere to this strict pattern. In an analysis of addressee facial contributions in dialogue, such as raising eyebrows to indicate surprise or grimacing to display a response appropriate for the speaker's described situation, Bavelas and colleagues ([Bavelas and Chovil, 1997](#); [Bavelas et al., 2002](#); [Bavelas and Gerwing, 2011](#)) demonstrated that this type of feedback did not fit the turn-coordinating definition of backchannels as occurring typically at transition relevance places. They go so far as to suggest, "the fact that addressees' facial contributions are usually simultaneous with the speaker's speech raises interesting questions about the utility and viability of the concept of 'turn taking'" ([Bavelas and Gerwing, 2011](#):190).

Our analysis and experiment will focus on function rather than placement. For both the collection of examples and the experimental stimuli, we use verbal generic and specific responses. The use of verbal responses, which typically do not

overlap with speakers' talk, allowed for a manipulation of the stimulus material in which a backchannel from one category was removed and replaced with another from the other category (for example where generic *yeah* occurred in the first condition, the specific *wow* occurred in the second). This manipulation allowed us to test the effect of categorical distinctions of backchannels within identical narrative contexts.

### 2.3. Backchannels looking forward

Within previous research, of both function and placement, the primary focus has been on relating a particular backchannel to the speaker's previous talk. In so doing, researchers have drawn attention away from the role backchannels play in the continuing, unfolding talk in a multi-turn utterance, or how the speaker might take feedback into account as they continue their talk. In contrast with these paradigms of research, a number of studies have demonstrated that speakers incorporate feedback into the development of their talk (Norrick, 2012).

By systematically altering which particular modalities were available to pairs engaged in a joint activity, researchers have shown that the types of backchannels an addressee provides, and where in the course of the speaker's talk these backchannels occur, is dependent on which communication channels are available in a particular context. Addressees who listened to a speaker on a telephone, for example, were much less likely to produce facial displays in response to speaker's talk, relying instead on verbal backchannels (Chovil, 1991). Similarly, in a study on tasks involving physical workspaces and strict roles with directors who did most of the speaking and matchers who provided feedback to directors, matchers provided feedback not through verbal backchannels but through their actions, and this feedback was likely to be initiated during the director's talk rather than at the end (Clark and Krych, 2004). Speakers engaged in this joint task actively monitored these actions as a means to incrementally adjust their speech as they would verbal backchannels in other tasks.

Experimental studies that have explored the role of backchannels in modulating speakers' moment-by-moment talk have typically done so by having pairs participate in highly structured tasks, involving clearly delineated goals and roles (e.g. Clark and Krych, 2004). This method increases control over the dialogue and reduces variation (Bavelas, 2005). There are indeed clear distinctions in the type of language used, including backchannels, depending on the situational context. Bangerter and Clark (2003), for example, compared task settings with informal conversation and found distinct patterns of use for different types of backchannels, including generics.

Within a collaborative account of language, all communication is a joint project in which the two interactants seek to accomplish the goal of social sharing together (Bavelas et al., 2000). This is true even of narrations, such as story telling or gossip, in which one speaker is likely to hold sole access to the information. Backchannels should therefore play a similar role in structuring the ongoing activity and influencing the speaker's talk as they do in explicitly task-oriented dialogue. Indeed, from a more global narrative perspective it has been shown that the type and quantity of backchannel tokens provided by the addressee influenced the structure and quality of speakers' narratives (Bavelas et al., 2000). Backchannels displaying uptake of the story content, such as markers of information state (*oh*; Heritage, 1984; Norrick, 2010b) and assessments (*wow*; Goodwin, 1986), are likely to be responded to explicitly by speakers in the directly subsequent talk (Norrick, 2010b). Yet to be explored however is how addressee backchannels systematically influence the unfolding structure of a narrative.

## 3. In-depth analysis of generic and specific backchannels in spontaneous dialogue

Following a perspective emphasizing the collaborative nature of dialogue as joint action, we analyzed how backchannels used in unstructured conversations shape the speaker's continuing talk in systematic patterns, focusing on discourse relationships between turns before and after target backchannels in an audio corpus of spontaneous dialogue. The corpus consists of conversations between pairs of undergraduate students at the University of California, Santa Cruz. Students participated in the collection of the conversations in return for course credit. Conversations were 12 min in length. These unstructured and loosely topical dialogues began with participants' discussing bad roommate experiences they may have had, but subsequent conversation was not controlled. Participants typically took turns telling stories of previous experiences with roommates, allowing for the collection of a variety of backchannels in the context of collaborative narration. Thirty conversations from this corpus were reviewed in total, with 20 one- to two-minute interactions selected for in-depth analysis, focusing on the moment-by-moment collaborative construction of the dialogue through the active and overt participation of both speakers and addressees.

### 3.1. Generic backchannels

As previous literature suggests, generic backchannels in this corpus were typically taken as displays of comprehension and continued attention. Across analyses, generic backchannels are viewed as indications that the previous talk has been received and comprehended, and are taken by speakers as permission to continue (Bangerter and Clark, 2003;

Goodwin, 1986; Schegloff, 1982). Importantly, after generic backchannels speakers continue in a systematic way. In task-oriented dialogue, in which projects are divided into a hierarchy of joint actions, generic backchannels are used in transitions from one subtask to another at the same level of the hierarchy (Bangerter and Clark, 2003). Similarly, Goodwin (1986) suggested that generic backchannels act as bridges between two units. In the context of casual conversation, which typically consists of smaller units of storytelling and narration, such as gossip, the units being bridged, or the subtasks of the joint activity, are expositions of discourse events. Thus, after generic backchannels, speakers continue their narrative by presenting new information. Typically this consists of presenting the next event of the narrative. The following transcript presents a single narration in which the addressee provided a series of generic backchannels.

- (4)
- 26 S5: Didn't Miss Lewis ever tell you about like her nephew or something  
 27 S6: ([no she did]) I probably forgot  
 28 S5: In the navy  
 29 S6: Mm[mm I dunno  
 30 S5: [ok- ok she had this nephew that was like- he was in the navy  
 31 and you have to be short cause to fit in the submarine you know  
 32 S6: Uh huh  
 33 S5: like cause they only make it like a certain height and he was like  
 34 only 5'6" or 8"  
 35 S6: Uh huh  
 36 S5: and then like he had a growth spurt while he was in the navy  
 37 S6: Uh huh  
 38 S5: and this is like bef- when he was twenty or twentyone and he  
 39 turned to like six something.

Prior to this point in the conversation, the two conversational participants were discussing whether girls or boys grow taller later in development. In order to argue her point that it is in fact males who go through their growth spurts later, S5 presents a narrative she learned from her teacher which she introduces in line 26 and begins in earnest in line 30. Following each *uh huh* from her addressee (lines 32, 35, and 37), S5 presents new information that develops the narrative along its current trajectory. For example after describing the height of the nephew beginning in line 34 as “only 5'6" or 8",” in line 36 she continues by presenting a discourse-new event, namely the nephew's growth spurt in the navy. This is responded to with another generic backchannel, after which S5 again presents discourse-new information, the age at which this occurred, confirming her proposal that males do indeed continue to grow taller later in adolescence.

With a proactive perspective on backchannels in dialogue, the two interlocutors can be seen as creating the discourse together. At each point in which new information is presented, the addressee accepted this information and displayed understanding through her use of backchannel communication. We argue that it was not necessarily S5's goal to construct a multi-turn utterance when she began her tale at line 30; rather, it was the joint process of presenting and accepting discourse events and relevant information, through generic backchannels, that lead to the construction of the speaker's narrative. After each generic backchannel, the speaker continued her story along a steady trajectory, building on the last event with a discourse-new event that did not attempt to redefine or embellish on the information presented in the last turn.

### 3.2. Specific backchannels

A different pattern is found with specific backchannels. Like generic backchannels, specific backchannels demonstrate continued attention. But specific backchannels also provide additional information, such as marking the speaker's talk as discourse-new or providing the addressees' affective response (Gardner, 2001). Rather than continue on with their stories, speakers take specific backchannels as cues for confirming the information of the previous turn in an elaborative or explanative manner. The following transcript presents a single narration in which the addressee provided a specific backchannels in line 45. In this conversation, a student, S7, is explaining to another the relationship between his girlfriend and her roommate. He then discusses a particular night in which he spent an evening at their shared apartment before returning to discussing their relationship status.

- (5)
- 40 S7: I look back and one of the dudes is following us right, like straight  
 41 walking behind us and I'm like look. And she turns around and is  
 42 like oh my god. As we start hella wa:lking and like hella tu:ring and  
 43 he's like following us follows us the whole way all the way um like

- 44 the stairs from the college 8 stop  
 45 S8: Oh my gosh  
 46 S7: A:ll the way down. So like- I had to spend the night last night  
 47 because she wouldn't let me walk back and like my knees messed  
 48 up and she was like hella scared and so I had to stay there. But like  
 49 I don't know like- like I don't know if- cuz they get along, like for  
 50 roommates but they're not best of friends, you know  
 51 S8: Yeah.  
 52 S7: she usually goes to sleep but then we'll be up and then we'll like  
 53 post for like twenty minutes or so and then we'll just start talking

The addressee provided two backchannels during the speaker's talk. As observed earlier with generic backchannels, in line 51, the addressee's generic *yeah* was responded to with a continuation of the story. The speaker mentions that the two roommates are not best friends in the prior line, and in the directly subsequent talk described a discourse-new event, in which the roommate attempted to sleep while he and his girlfriend talked.

In contrast, the specific backchannel in line 45, *oh my gosh*, was produced as an affective or informational response to the prior utterance. The addressee is responding to the content of the speaker's talk, that they were being followed, rather than simply acknowledging continued attention or comprehension. Unlike responses to generic backchannels, in which speakers continue on with their stories, with specific backchannels, speakers continue with an elaboration of the content to which the backchannel responded. In this case, in the utterance prior to the backchannel, the speaker described being followed and where, emphasizing and clarifying the information in the following turn with "all the way down."

Importantly, the responses to elaborations are distinct from patterns associated with other-initiated repair (Schegloff, 1997). By providing affective responses to the content of the previous turn, addressees are not indicating any trouble in comprehension of the talk. While similar to other-initiated repair in that the specific backchannels are responded to with information about the event presented in the last turn, the difference is that the information is not meant to be a reiteration to aid comprehension. Rather, the elaborative next turns provide discourse-new information about the same discourse event, as a means of implicitly accepting the addressee's stance presented in their backchannel. Indeed, in contrast with the pattern found in responding to generic backchannels, when addressees provided specific backchannels, speakers often explicitly commented on addressees' responses in their next turn.

- (6)  
 54 S9: and then he had like jumper cables right? And that was like what  
 55 we needed, and then he did it, and I guess he did it wrong or  
 56 something  
 57 S10: o:h  
 58 S9: cause it cause it just like messed up the car  
 59 S10: oh my go:d sca:ry  
 60 S9: Yeah! So then, whe:n... the car started, but the thing is that um,  
 61 the... It wouldn't accelerate

In the previous transcript, the storyteller, S9, is describing an event in which her car had broken down, stranding her and her friends. Following the previous analysis above, S9 provides an elaborative next turn following the *oh* produced in response to "I guess he did it wrong or something." In line 59, the addressee, S10, provides a second specific backchannel, "oh my god scary." In response to this, the storyteller provides talk indicative of the relationship between specific backchannels and elaborative next utterances. First, she provides an explicit turn-initial uptake of the addressee's response, *Yeah!* Explicit acknowledgement of the backchannel was only found following specific backchannels in the conversations analyzed, indicating that they do indeed function distinctly from generic continuers in the storytelling activity. The *Yeah!* is followed by a false start: What appears to be a continuation, starting with the discourse markers *so then*, is abandoned and elaborative information on the *messed up*-edness of the car is provided. These features of the speaker's next talk support an analysis of elaborative next turns as the sequentially preferred response following specific backchannels, with continuations of the narrative following generic backchannels.

From the above analysis of backchannels in spontaneously produced narrative dialogues, we derived two hypotheses about the relationship between generic and specific backchannels and the unfolding discourse. We propose that for each of two categories of backchannel, generic or specific, different types of speaker talk will be more likely to follow. After generic backchannels, the next utterance by the speaker is more likely to continue the narrative, with the speaker introducing some next event or other material that is new to the discourse. In contrast, after specific backchannels, the next utterance by the speaker is more likely to elaborate on the information to which the specific backchannel responded.

Although addressees' shaping storytelling is not surprising given the collaborative nature of the joint action of dialogue, this is the first analysis to demonstrate that backchannels influence not just the global level of story quality (e.g. Bavelas et al., 2000), but a speaker's turn-by-turn narrative development.

#### 4. Experiment: conversation completion

The hypotheses derived from the qualitative analysis of the storytelling corpus were tested using an experimental story completion paradigm. In the experiment presented below, we provided participants with transcripts of storytelling interactions collected from the same corpus used in the analysis above, up to a critical target backchannel, and asked them to make up the next turn in the conversation in two counter-balanced conditions. Participants read these short conversations up to a generic or specific backchannel, and then placed themselves in the role of the speaker, or storyteller, and provided what they thought would be an example of what this speaker would likely say next. If the pattern of responses following different categories of backchannels varies as proposed, participants should be sensitive to this relationship and correspondingly vary the next turns they provide. Across the two conditions of the experiment, transcripts were matched in all respects except the critical response, which was either a generic or specific backchannel, controlling for any effect the propositional content of the speaker's turn may have had on how the stories were continued.

##### 4.1. Methodology

Twenty interactions were transcribed, 10 with naturally occurring generic backchannels, either *mhm*, *uh huh*, or *yeah*, and 10 with naturally occurring specific backchannels, either *oh*, *really*, *wow*, or *whoa*. From these 20 stimuli, an additional 20 were created with the naturally occurring backchannel replaced with a backchannel from the opposite category. Two lists were created, each containing equal numbers of natural and altered stimuli and equal numbers of generic and specific backchannels in the crucial location. The order of presentation was the same across lists and was pseudo-randomized so that the backchannels of one category did not follow one another.

Sixty students from the University of California Santa Cruz (42 female) participated in exchange for course credit. Participants were randomly assigned to one list, and so only read one version of each stimulus. Participants were tested individually, with each session lasting about 30 min. Each list consisted of 20 short dialogues, presented in play dialogue format with randomized, gender-balanced names. After each dialogue, which ended with either a specific or generic backchannel, participants were given the name of the storyteller and a space to write in what they think would be a likely next line of talk.

Two raters trained to distinguish between discourse continuations and elaborations judged each example next turn, categorizing the relationship between the last speaker's talk and the participant's proposed subsequent talk as either a discourse continuation, elaboration, or neither. Continuations were any next turn that provided some new event in the narrative, whereas elaborations were any next turn that provided additional information of the same discourse event that was the focus of the speaker's turn prior to the critical backchannel. Additional information included explanations, elaborations, and re-interpretations. The neither category was used for any response that was not relevant to the development of the story, and included questions such as *so what about you?* as well as one-word responses such as *Yup*. Where participants provided two sentence answers, raters coded only the first. Raters were blind to hypotheses and were not provided the target backchannels. The inter-rater reliability for the two trained raters was Kappa = .51,  $p < .001$ , a moderate agreement. All disagreements on coding were resolved jointly by the two raters and one of the authors. The raters also coded the participants' responses for the presence of turn-initial discourse markers.

##### 4.2. Results and discussion

Participant next turn responses were coded as continuations, elaborations, or neither (see Fig. 1 for an example stimulus with responses from each category). Data from participants who had 25% or more task-irrelevant *neither* responses were removed from analysis (a total of 4 participants). All other *neither* responses were dropped from the analysis (49 responses, 4% of total responses). Percentage of discourse continuation responses after generic and specific backchannels for the remaining participants were calculated, and these were tested as a within-subjects factor. More continuations followed generic backchannels ( $M = 36\%$ ,  $SD = 18\%$ ) than specific backchannels ( $M = 29\%$ ,  $SD = 20\%$ ), mean difference = 6.66%,  $t(55) = 2.27$ ,  $p = .028$  (see Fig. 2).

Inspection of the proposed next turns revealed that many of the suggested dialogue developments began with pragmatic devices. Under the conceptualization of three-part grounding sequences proposed by Clark and Shaefer (1987), turn-initial pragmatic devices display the speakers' understanding of mutual acceptance of the speakers' prior contribution. We analyzed the use of turn-initial acknowledgement tokens such as *yeah* separately from turn-initial

**Stimulus:**

- Steven:** We had one incident where the room like, smelled pretty bad when we got in there one time.
- David:** Haha.
- Steven:** I don't know, it was kind of an awkward smell but...
- David:** Dude, after um, after this weekend our whole house, like I woke up Monday morning the whole house just smelled like stale beer.
- Steven:** No way.
- David:** Our carpet is just like, so nasty. We're gonna get like one of those steam cleaners I think.
- Steven:** [*Mhm/Oh*]
- David:**

**Responses:***Continuation*

- I think I have to make some ground rules with my housemates.
- The bathroom was disgusting too. We'll probably need more than just the steamer.

*Elaboration*

- Yeah, I heard steam cleaners are like the most effective way to clean carpets.
- They're expensive though.
- I feel like it's the only way to get the stains out, y'know?

*Neither*

- Haha, yup.

Fig. 1. Example transcript stimulus with generic backchannel target *mhm* and specific backchannel target *oh* presented across conditions. Participant responses demonstrate next turns coded as continuations, elaborations, and neither.

discourse markers, which included *so*, *and*, *well*, and *but* in our data. A greater number of acknowledgement tokens followed specific backchannels ( $M = 5.16$   $SD = 2.7$ ) than generic backchannels ( $M = 1.93$ ,  $SD = 2.0$ ),  $t(55) = 12.26$ ,  $p < .001$  (see Fig. 3). In contrast, analysis of the use of turn-initial discourse markers revealed a greater number of discourse markers following generic backchannels ( $M = 2.46$   $SD = 2.1$ ) than following specific backchannels ( $M = 1.23$ ,  $SD = 1.44$ ),  $t(55) = 5.30$ ,  $p < .001$ .

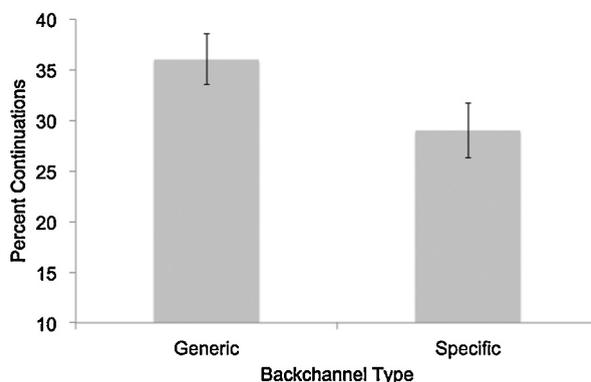


Fig. 2. Percentage of continuations after generic and specific backchannels, calculated as number of continuations divided by total number of continuations and elaborations.

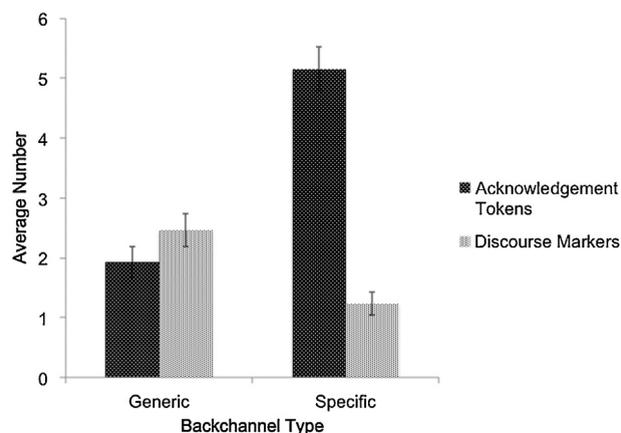


Fig. 3. Number of turn-initial pragmatic devices after generic and specific backchannels.

To summarize, participants provided a greater proportion of continuations following a generic backchannel than a specific backchannel, matching the predicted relationship between addressee response and the subsequent unfolding of the narrative discourse. More specifically, when participants read transcripts of conversation up to a generic backchannel, they were more likely to write a next turn that introduced some discourse-new event compared to the same transcript with a specific backchannel. After specific backchannels, participants provided proportionally fewer continuations and more elaborations of the speaker talk on which the specific backchannel commented.

Overall, elaborative next turns were a more common response. We believe that providing elaborations is easier than creatively suggesting how a story might develop. That is, less effort and thought is needed to elaborate on information present in the transcript as compared to inventing new information. This is particularly true given that participants had very limited access to the content of the stories, in some cases as little as three turns. However, even with increased effort toward creating continuations and limited access to the content of the stories, the type of backchannel provided by the addressee still influenced what participants thought would happen next in the development of the discourse. That is, even when they could write whatever they wanted as subsequent turns, and even when they may have been less able or motivated to create continuations, participants' suggested next turns were influenced by the addressee backchannel that they read.

Similarly, acknowledgement tokens were more common than discourse markers at the beginning of participants' responses. However, the prevalence of both acknowledgement tokens and discourse markers depended on the type of prior backchannel read. More acknowledgement tokens were used after specific than generic backchannels, and conversely more discourse markers were used after generic than specific backchannels. These contrastive effects demonstrate that participants treated the backchannels as distinct contributions to the dialogue, requiring distinct responses, further highlighting the functional distinction between generic and specific backchannels in shaping the content of directly subsequent speaker turns.

## 5. General discussion

If backchannels simply provided different types of responses to the speaker's multi-turn utterances, either as secondary signals in a conversation or as reactions to prior speech, there should be no systematic differences in either naturally-occurring or participant-proposed developments of narratives based on whether a specific or generic backchannel was used. However, in both an in-depth analysis of backchannels produced in a spontaneous face-to-face corpus and in a story-completion experiment, we found that backchannel addressee responses were proactive, shaping the unfolding narrative moment-by-moment.

Both our analysis and the conditional distinctions within the experiment relied on previously developed functional categories of backchannels, those of generic and specific responses (Goodwin, 1986; Schegloff, 1982; Stivers, 2008). Previously the categorical distinction between backchannel type has been based on either placement within speaker talk (Goodwin, 1986) or type of response displayed (Brunner, 1979; Gardner, 2001). Building on previous literature demonstrating a variety of possible speaker responses to addressee backchannels (Norrick, 2010b, 2012), we add to the distinction between generic and specific backchannels evidence of a systematic difference in how these categories steer unfolding narrative discourse.

The proactive nature of backchannel communication caused our participants to provide different examples as to what the speaker would most likely say next depending on the backchannel used. If they read a specific backchannel, including

tokens related to informational state of the addressee such as *oh* and *really*, as well as assessment tokens such as *wow* and *gee*, participants were more likely to provide an elaboration of the speaker's last turn. They were also more likely to begin their proposed talk with an acknowledgement token such as *yeah*, explicitly acknowledging the addressee's specific backchannel, replicating previous findings of speaker acknowledgements to addressee contributions in spontaneous talk (Norrick, 2010b, 2012). If they read a generic backchannel, participants were more likely to provide a continuation of the speaker's last turn. They were also more likely to begin their subsequent talk with a discourse marker such as *so*, *and*, *but*, or *well*, explicitly marking how the next event in the story should be interpreted with respect to the prior event. For example, *so* indicates the following talk is not contingent on the preceding talk (Bolden, 2009) and *well* indicates that upcoming information is relevant despite seeming as if it is not (Blakemore, 2002). So the type of backchannel affected both the content of the subsequent proposed turn and the choice of turn-initial pragmatic device. Acknowledgement tokens usefully recognize a specific backchannel's invitation to elaborate, and discourse markers usefully highlight how the next story event relates to the prior after a generic backchannel's suggestion for a continuation.

Previous researchers have demonstrated the role of backchannels in narrative dialogues (Bavelas et al., 2000; Beukeboom, 2009). They systematically modulated the way addressees responded to speakers, either through distraction or through the use of confederates instructed to behave in certain ways. Stories told to distracted addressees, while not different in length, were judged to be lower in quality with worse endings (Bavelas et al., 2000). The researchers suggested that undistracted addressees helped speakers finish stories smoothly and effectively. But distraction also drastically reduced the number of specific backchannels addressees provided (Bavelas et al., 2000). Based on the results of the present analysis we suggest that by not providing specific backchannels, distracted addressees likely reduced the amount of elaborative information that otherwise would have developed the narratives' more relevant or interesting features, which in turn would have contributed to more climactic endings. Addressees providing only generic backchannels left their speakers to simply tell their stories as a series of events, without highlighting important elements. This may have led to the repetitions and awkward justifications observed by Bavelas et al. (2000).

Our participants, while not active participants in the conversations, still made systematic predictions as to what was likely to be spoken next, given either a generic or specific backchannel from the addressee. This suggests that in the comprehension of dialogues, overhearers (or readers of dialogue text) may make use of predictive relationships across speaker and addressee contributions. A number of models of language comprehension have focused on the role of prediction (see e.g. Pickering and Garrod, 2013). Much work in language comprehension has focused on passive listeners presented with monologues; however, comprehension of dialogues likely involves predictions as well. For example, participants who listened to only half of a dialogue, a *halfalogue*, were more distracted than those who listened to a full dialogue (Emberson et al., 2010). Emberson et al. suggested that it was the reduction of predictability when the talk from only one interactant was available that lead to this increased distraction. Similarly, it is possible that overhearers listening to narrative dialogues (or readers of dialogue text) may make use of the proactive nature of backchannels as a means of predicting the type of information likely to be presented next, leading to faster discourse comprehension.

Future studies can expand on the findings presented here. In order to manipulate the target backchannel with texts, we were limited to verbal backchannels. In addition, the written modality left open the possibility that participants may have shaded the backchannels with additional prosodic or paralinguistic information in their readings of the dialogue. *Yeah*, as described above (section 2.1), could take on the function of a specific backchannel given a particular pronunciation and emphasis. While the punctuation of the transcription attempted to avoid such dramatic interpretations of generic backchannels, replicating results with audio stimuli would bolster the current findings. Furthermore, in focusing on spontaneous, loosely topical face-to-face conversations, our findings are also limited to the specific context of story telling. In their exploration of task-oriented discourse, Bangerter and Clark (2003) found almost no specific backchannels. Instead, Bangerter and Clark contrasted generic backchannels with tokens of agreement and consent such as *right* and *okay*. Similarly, in an analysis of the role of listener feedback on storytelling, Norrick (2010b) distinguished between assessments and information state responses, categories we combined as context specific backchannels. Further exploration will be needed to test the collaborative, predictive relationships between different types of backchannels and the unfolding speaker talk.

The story completion task is a promising method for future research. We suspect that effects on original conversational participants may be larger than those reported here for strangers reading conversational text. As discussed earlier, experimental participants may have had an easier time coming up with elaborations than continuations. This could potentially shrink the size of the effect. For overhearers or over-readers, longer turns, which provide more narrative content, might produce fewer elaborations overall, and thus increase the distinction between specific and generic backchannels and their effects on subsequent speaker talk.

## 6. Conclusion

Together, the analysis of generic and specific backchannels and the experiment on the relationship between these backchannels and the subsequent development of a speaker's narrative demonstrate the proactive nature of

backchannels in the collaborative context of storytelling. When overhearers suggested likely next developments of the narrative, their responses to context-generic continuers such as *mhm* and *uhuh* differed from their responses to context-specific assessments such as *oh* and *wow*. When they read generic backchannels they were more likely to continue the story with some next event. When they read specific backchannels they were more likely to elaborate on the information provided before the backchannel.

The novel finding of a regularity in the expected talk after different types of backchannels builds on previous research on how addressees co-construct talk. Addressees collaborate directly in the moment-by-moment creation of talk, even in the context of narrative, where the speaker likely holds strong if not singular epistemic access to the content. More broadly, we suggest that a fully developed theory of backchannel communication includes not only their function as responses to speaker talk, but also their role in pushing unfolding speakers' talk along particular trajectories.

## Acknowledgements

This research was supported by faculty research funds granted by the University of California, Santa Cruz. Funding for Open Access was provided by the University of California, Santa Cruz, Open Access Fund. We thank our research assistants who aided in data collection and coding, with a special thanks to Jasper Hall, Christopher Maniotes, and Heather Bach. Funding for Open Access provided by the University of California, Santa Cruz, Open Access Fund.

## References

- Bangerter, Adrian, Clark, Herbert H., 2003. Navigating joint projects with dialogue. *Cogn. Sci.* 27, 195–225.
- Bavelas, Janet, 2005. The two solitudes: reconciling social psychology and language and social interaction. In: Fitch, K., Sanders, R. (Eds.), *Handbook of Language and Social Interaction*. Erlbaum, Mahwah, NJ, pp. 179–200.
- Bavelas, Janet B., Chovil, Nicole, 1997. Faces in dialogue. In: Russell, J.A., Fernandez-Dols, J.M. (Eds.), *The Psychology of Facial Expression*. Cambridge University Press, Cambridge, England, pp. 334–346.
- Bavelas, Janet B., Gerwing, Jennifer, 2011. The listener as addressee in face-to-face dialogue. *Int. J. Listening* 25, 178–198.
- Bavelas, Janet B., Coates, Linda, Johnson, Trudy, 2000. Listeners as co-narrators. *J. Pers. Soc. Psychol.* 79, 941–952.
- Bavelas, Janet B., Coates, Linda, Johnson, Trudy, 2002. Listener responses as a collaborative process: the role of gaze. *J. Commun.* 52, 566–580.
- Bertrand, Roxane, Ferré, Gaelle, Blache, Phillippe, Espesser, Robert, Rauzy, Stéphane, 2007. Backchannels revisited from a multimodal perspective. In: *Proceedings of Auditory-Visual Speech Processing*, Hilvarenbeek, Netherlands.
- Beukeboom, Camiel, 2009. When words feel right: how affective expressions of listeners change a speaker's language use. *Eur. J. Soc. Psychol.* 39, 747–756.
- Blakemore, Diane, 2002. *Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers*. Cambridge University Press, Cambridge.
- Bolden, Galina, 2009. Implementing incipient actions: the discourse marker 'so' in English conversation. *J. Pragmatics* 41, 974–998.
- Bruner, Lawrence, 1979. Smiles can be back channels. *J. Pers. Soc. Psychol.* 37 (5), 728–734.
- Chovil, Nicole, 1991. Social determinants of facial displays. *J. Nonverbal Behav.* 15, 141–153.
- Clancy, Patricia, Thompson, Sandra, Suzuki, Ryoko, Tao, Hongyin, 1996. The conversational use of reaction tokens in English, Japanese, and Mandarin. *J. Pragmatics* 26, 355–387.
- Clark, Herbert H., 1996. *Using Language*. Cambridge University Press, New York.
- Clark, Herbert H., Krych, Meredith A., 2004. Speaking while monitoring addressees for understanding. *J. Mem. Lang.* 50, 62–81.
- Clark, Herbert H., Murphy, Gregory L., 1982. Audience design in meaning and reference. *Adv. Psychol.* 9, 287–299.
- Clark, Herbert H., Shaefer, Edward F., 1987. Collaborating on contributions to conversations. *Lang. Cogn. Process.* 2, 19–41.
- Dittmann, Allen, Llewellyn, Lynn, 1986. Relationship between vocalizations and head nods as listener responses. *J. Pers. Soc. Psychol.* 9, 79–84.
- Duncan Jr., Starkey, 1972. Some signals and rules for taking speaking turns in conversations. *J. Pers. Soc. Psychol.* 23, 283–292.
- Duncan Jr., Starkey, 1974. On the structure of speaker–auditor interaction during speaking turns. *Lang. Soc.* 2, 161–180.
- Duncan Jr., Starkey, Fiske, Donald W., 1977. *Face-to-face Interaction: Research, Methods, and Theory*. Wiley, New York.
- Emberson, Lauren, Lupyán, Gary, Goldstein, Michael H., Spivey, Michael J., 2010. Overheard cell-phone conversations: when less speech is more distracting. *Psychol. Sci.* 21, 1383–1388.
- Fox Tree, Jean E., 1999. Listening in on monologues and dialogues. *Discourse Process.* 27, 35–53.
- Fox Tree, Jean E., Schrock, Josef C., 2002. Basic meanings of *you know* and *I mean*. *J. Pragmatics* 34, 727–747.
- Fries, Charles C., 1952. *The Structure of English: An Introduction to the Construction of English Sentences*. Brace, Harcourt.
- Gardner, Rod, 2001. *When Listeners Talk*. Benjamins, Amsterdam.
- Goodwin, Charles, 1986. Between and within: alternative sequential treatments of continuers and assessments. *Hum. Stud.* 9, 205–217.
- Heritage, John, 1984. A change-of-state token and aspects of its sequential placement. In: Atkinson, J. Maxwell, Heritage, John (Eds.), *Structures of Social Action*. Cambridge University Press, Cambridge, pp. 299–347.
- Kendon, Adam, 1967. Some functions of gaze direction in social interaction. *Acta Psychol. (Amst.)* 26, 22–63.
- Koiso, Hanae, Horiuchi, Yasuo, Tutiya, Syun, Ichikawa, Akira, Den, Yasuharu, 1998. An analysis of turn-taking and backchannels based on prosodic and syntactic features in Japanese Map Task dialogs. *Lang. Speech* 41 (3–4), 295–321.
- Morency, Louis P., de Kok, Iwan, Gratch, Jonathan, 2010. A probabilistic multimodal approach for predicting listener backchannels. *Auton. Agent Multi-Agent Syst.* 20, 70–84.

- Norrick, Neal R., 2010a. [Listening practices in television interviews](#). *J. Pragmatics* 42, 525–543.
- Norrick, Neal R., 2010b. [Incorporating listener evaluation into stories](#). *Narrative Inq.* 20, 183–204.
- Norrick, Neal R., 2012. [Listening practices in English conversation: the responses responses elicit](#). *J. Pragmatics* 44, 566–576.
- O'Connell, Daniel C., Kowal, Sabine, Kaltenbacher, Erika, 1990. [Turn-taking: a critical analysis of the research tradition](#). *J. Psycholinguist. Res.* 19, 345–373.
- Pickering, Martin, Garrod, Simon, 2013. [An integrated theory of language production and comprehension](#). *Behav. Brain Sci.* 36, 329–392.
- Sacks, Harvey, Schegloff, Emanuel A., Jefferson, Gail, 1974. [A simplest systematics for the organization of turn-taking in conversation](#). *Language* 50, 696–735.
- Schegloff, Emanuel A., 1982. [Discourse as an interactional achievement: some uses of 'uh huh' and other things that come between sentences](#). In: Tannen, D. (Ed.), *Analyzing Discourse: Text and Talk*. Georgetown University Press, Washington, DC.
- Schegloff, Emanuel A., 1997. [Practices and actions: boundary cases of other-initiated repair](#). *Discourse Process.* 23, 499–545.
- Stivers, Tanya, 2008. [Stance, alignment and affiliation during storytelling: when nodding is a token of affiliation](#). *Res. Lang. Soc. Interact.* 41 (1), 31–57.
- Tomlinson, Jack, Fox Tree, Jean E., 2011. [Listeners' comprehension of uptalk in spontaneous speech](#). *Cognition* 119, 58–69.
- Ward, Nigel, Tsukahara, Wataru, 2000. [Prosodic features which cue back-channel responses in English and Japanese](#). *J. Pragmatics* 32, 1177–1207.
- Yngve, Victor H., 1970. [On getting a word in edgewise](#). In: Campbell, M.A. (Ed.), *Papers from the Sixth Regional Meeting, Chicago Linguistics Society*. Department of Linguistics, University of Chicago, Chicago, pp. 567–578.