

The Role of the Current Speaker in Conversational Turn Taking – Theoretical, Experimental, and Corpus Linguistic Perspectives on Speaker Contributions to Aligned Turn-Timing

Workshop @HU Berlin, 14/15 January 2021 (Central European Time [UTC +1h])

<https://www.angl.hu-berlin.de/news/conferences/workshop-current-speaker>

Day 1

Day 2

09:00 – 09:15 Welcome and Registration / Welcome Back	
<p>Session 1</p> <p>09:15 – 09:45 Mathias Barthel – Introduction</p> <p>09:45 – 11:00 Invited Talk: Judith Holler, MPI for Psycholinguistics – Turn-timing and the body</p>	<p>09:15 – 10:30 Invited Talk: Lorenza Mondada, University of Basel – Ongoing adjustments between current and next: early projections and anticipations in talk and embodied interaction</p> <p>10:30 – 11:00 Nele Pöldvere, Victoria Johansson & Carita Paradis – Social and cognitive explanations of dialogic resonance in everyday conversation</p>
11:00 – 11:15 Coffee Break	
<p>Session 2</p> <p>11:15 – 11:45 Jaroslav R. Lelonkiewicz & Chiara Gambi – Making oneself predictable in linguistic interactions</p> <p>11:45 – 12:15 Christine Howes & Mary Lavelle – Mis-taking turns: Conversing with a patient with schizophrenia</p> <p>12:15 – 12:45 Anna Katharina Kuhlen & Rasha Abdel Rahman – Mental Chronometry of Speaking in Dialogue: Semantic Interference Swings to Facilitation Across Speaker Turns</p>	<p>11:15 – 11:45 Marcin Włodarczak & Mattias Heldner – Turn-taking in conversation from the larynx down</p> <p>11:45 – 12:15 Marina Noelia Cantarutti – When two are “doing being” one: Co-animation and association in collaborative turn sequences</p> <p>12:15 – 12:45 Christoph Rühlemann & Stephan Gries – Turn-final lengthening is really rallentando across (much of) the turn</p>
12:45 – 14:00 Lunch Break (13:30 – 14:00 Gather Town Lobby)	
<p>Session 3</p> <p>14:00 – 14:45 Special Guest</p> <p>14:45 – 15:15 Laurel Brehm – How we take turns: A new statistical modeling approach</p>	<p>14:00 – 14:30 Viktória Horváth, Valéria Krepsz, Dorottya Gyarmathy, Anna Huszár & Ágnes Hátori – Dynamic changes of speech patterns as cues of smooth turn-takings in Hungarian triadic conversations</p> <p>14:30 – 15:45 Invited Talk: Peter Auer, Albert-Ludwigs-University of Freiburg – Gaze selects the next speaker in answers to German <i>ihr</i>-questions</p>
15:15 – 15:45 Coffee Break	
<p>Session 4</p> <p>15:45 – 17:00 Invited Talk: JP de Ruiter, Tufts University – The role of syntactic and prosodic cues in the anticipation of conversational turns</p>	<p>15:45 – 16:00 Coffee Break</p> <p>16:00 – 17:00 Panel Discussion on Methodological Fusion in Interaction Research</p>
(17:00 – 18:00 Gather Town Lobby)	

Making oneself predictable in linguistic interactions

Jarosław R. Lelonkiewicz & Chiara Gambi

While language production is a highly demanding task, conversational partners are known to coordinate their turns with striking precision. Among the mechanisms that allow them to do so is listeners' ability to predict what the speaker will say, and thus to prepare their response in advance. But do speakers also play a role in this process? We hypothesised that speakers contribute by using coordination smoothers – in particular by making their turns easier to predict. To test this, we asked participants to type definitions for common English words, either on their own ($n = 26$ individuals) or interacting with a partner ($n = 18$ pairs), and we measured the timing with which they produced the definitions. In a post-test, additional participants ($n = 55$) attempted to predict the final word of these definitions. We found that interacting speakers initiated their turns with less variable delays than solo individuals. In contrast, our post-test measures suggested that jointly produced definitions were in fact of lower predictability than those produced by individuals, but the analysis revealed these findings were likely confounded by task difficulty. We propose that the reduction in temporal variability observed for interacting speakers may facilitate prediction and thus act as a coordination smoother in linguistic interactions.

Mis-taking turns: Conversing with a patient with schizophrenia

Christine Howes
University of Gothenburg
christine.howes@gu.se

Mary Lavelle
City University, London
mary.lavelle@city.ac.uk

The *collaborative model* of communication states that conversation is a collaborative process (Clark 1996) where speakers and listeners produce information together, continuously coordinating and collaborating to establish each other's understanding and incrementally co-constructing the evolving content. Smooth turn exchange is achieved through tight coordination of interlocutors' verbal and non-verbal communication (Bavelas et al., 2002). Turn exchange becomes problematic when this tightly coordinated communication deviates from expectations. Under these circumstances, turn exchange processes, and the strategies that people use to overcome them, become more overt.

Patients with a diagnosis of schizophrenia are one of the most socially excluded in society (Addington & Addington, 2008). A central and debilitating feature of schizophrenia, which may contribute to patients' social exclusion, is patients' difficulty interacting with others, including the ability to 'mesh' their turns appropriately (Mueser et al., 1991). Interactions involving patients with a diagnosis of schizophrenia therefore offer an opportunity to observe the strategies that people employ when turn exchange is problematic, and shed light on how 'normal' turn exchanges are managed.

We use data from a corpus of triadic conversations originally collected to investigate only non-verbal behaviour (Lavelle et al., 2013). The corpus contains 20 dialogues involving one patient with a diagnosis of schizophrenia and two healthy controls (unaware of the patient's diagnosis) and 20 dialogues involving three healthy participants. We hypothesise that turn exchange will not be as smooth in dialogues including a patient with schizophrenia. This should be evident both in patients' lack of uptake of the turn when their interacting partners expect it and uncertainty regarding the cues speakers give in indicating that they have finished their turn or in selecting the next speaker (Schegloff et al., 1974).

Results show greater within turn pauses in patients' partners, when they leave space for a turn exchange that the patient does not take up (Howes et al., 2017) as well as greater delay before the patients' partners take the floor (see figure 1). Following McCabe & Lavelle (2012), we identify the increasingly explicit strategies that participants employ in offering the floor to the patient from nonverbal cues to explicit direct questioning. Initially, speakers invite the patient to take their turn through use of gaze and pauses, which the patient may decline by avoiding speaker gaze. Following failed attempts to offer the floor nonverbally to the patient, we see participants employ more explicit strategies such as verbal invitations for the patient to speak (e.g. 'what do you think?'), which are more likely to result in patients' taking their turn, but are rarely seen in the control group dialogues.

This shows how, despite the lack of clarity about who should take the floor in patient interactions, and patients' problematic floor change cues, their interlocutors can nevertheless adapt the strategies they use to manage the coordination of smooth turn exchanges.

References

- Addington J. and Addington D. (2008) Social and cognitive functioning in psychosis. *Schizophrenia Research* 99, 176–181.
- Clark (1996). *Using Language*. Cambridge University Press.
- Howes, C., Lavelle, M., Healey, P. G. T., Hough, J., and McCabe, R. (2017). Disfluencies in dialogues with patients with schizophrenia. In *Proceedings of the 39th Annual Meeting of the Cognitive Science Society*, London, UK.
- Lavelle, M., Healey, P. G., and McCabe, R. (2013). Is nonverbal communication disrupted in interactions involving patients with schizophrenia? *Schizophrenia Bulletin*, 39(5), 1150–1158.
- McCabe, R., and Lavelle, M. (2012) Participation in multiparty interaction: A single case analysis of a patient with schizophrenia. In *Proceedings of Video Analysis on Work, Interaction and Technology: A Festschrift for Christian Heath*, Bayreuth, Landratesaal, 97-108.
- Mueser, K. T., Bellack, A. S., Douglas, M. S., & Morrison, R. L. (1991). Prevalence and stability of social skill deficits in schizophrenia. *Schizophrenia Research*, 5(2), 167–176.
- Sacks, H., Schegloff, E., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696–735.

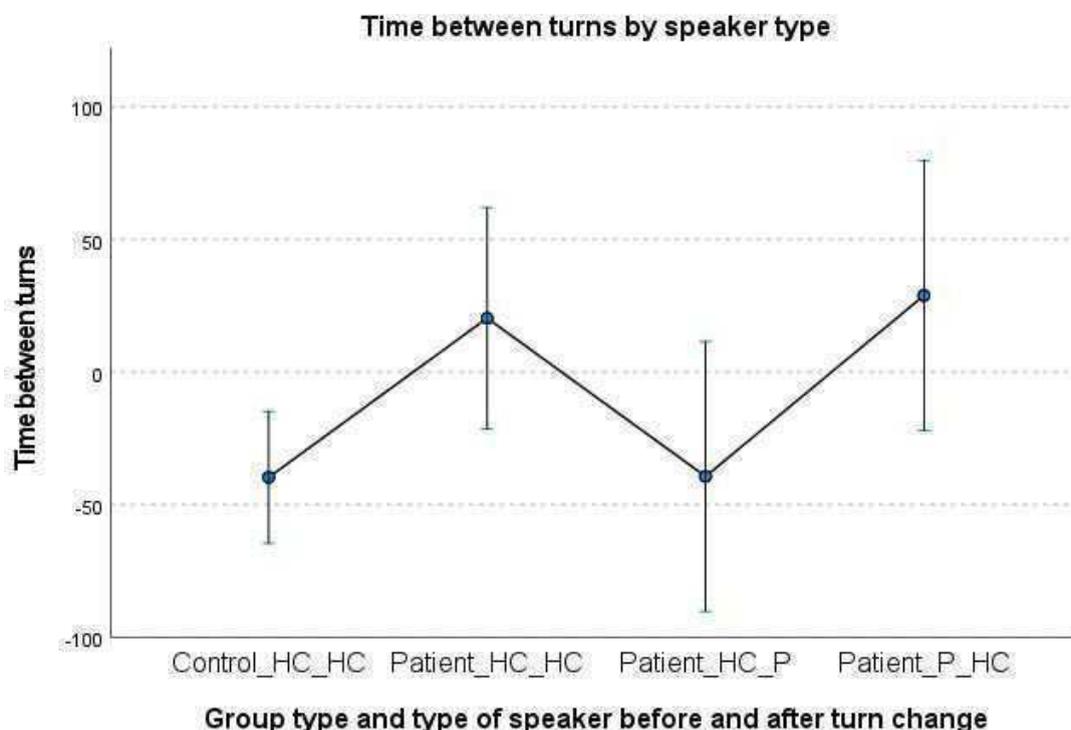


Figure 1: Time between speaker turns in milliseconds by group type and speaker and next speaker type (HC = Healthy Control; P = Patient)

Mental Chronometry of Speaking in Dialogue: Semantic Interference Swings to Facilitation Across Speaker Turns

Anna K. Kuhlen & Rasha Abdel Rahman

Department of Psychology, Humboldt-Universität zu Berlin, Berlin, Germany

Numerous studies demonstrate that the production of words is delayed when speakers process in close temporal proximity semantically related words. These studies are based on experimental settings in which isolated speakers name pictures alongside semantically related or unrelated context words. These settings are quite removed from those under which we typically speak. Specifically, in spontaneously occurring conversations interlocutors' utterances are frequently semantically related to each other. This study demonstrates that semantic interference disappears, and can even turn into facilitation, when semantically related words are embedded in a collaborative turn sequence.

Experiment 1 and 3 (each N=32 university students) implemented a picture-word interference task in a game played between two participants during which one named the distractor word and, after a stimulus-onset-asynchrony (SOA) of -100ms or -650ms, the other named a semantically related or unrelated target picture. Although picture-word interference is typically quite robust, in this communicative setting we did not observe interference. Instead, in Experiment 3, which encouraged a focus on the conceptual relationship between distractor word and target picture, we observed semantic facilitation at SOA-650. Semantic interference reappeared with identical experimental parameters in a traditional single-subject picture-word interference setting (Experiment 2, N=32).

We conclude that the processes leading to semantic interference in single-subject settings are attenuated in a dual-subject communicative turn sequence. Depending on communicative context and inter-turn interval semantic relatedness between the partner's and the own utterances does not interfere and may even facilitate speech production. We suggest that a communicative setting promotes processing the conceptual relationship between distractor and target and hence enhances semantic priming. Our findings carry implications for smooth turn transitions and highlight the importance of investigating language production in settings in which it typically occurs, namely in social interaction.

The study's hypotheses, as well as its experimental procedures and analyses were pre-registered.

How we take turns: A new statistical modeling approach

Laurel Brehm

MPI for Psycholinguistics

As measured by gaps between turns, speakers vary in the success of their coordination in conversation. In many corpora, the mean turn gap between speakers is about 200 ms (see e.g. Stivers, et al, 2009; Levinson & Torreira, 2015), but this single number hides a large range. Turn gap can be negative, reflecting overlaps between talkers, or can be quite long, leaving dead air. One important question examined across the literature is why these gaps arise. Existing experimental work suggests that turn gaps could be due to delays in planning speech or imprecision in identifying prior turn end in preparation for timely launching of speech (e.g. Corps, Crossley, Gambi, & Pickering, 2018; Magyari & de Ruyter, 2012). I focus on the question of timely launching in the current work. What do speakers need to know about the prior utterance in order to launch a pre-planned message with limited turn gap?

In order to assess the factors that contribute to launching timely speech, I propose a statistical solution that can be used to examine both experimental and corpus data: build and evaluate statistical models that allow the researcher to quantify not only how much a predictor increases or decreases turn gap, but also to establish which predictors are independent and to rank which are most important. This solution is model selection by k-fold cross validation: a simple machine-learning technique that is a conceptual extension of regression modeling (see e.g. Arlot & Celisse, 2010; Zhang & Yang, 2015 for review).

In this talk, I describe how cross validation models can be used to build and evaluate models of speakers' turn gap with a potentially unlimited set of predictors, how this allows us to assess which predictors are most important, and how this shows which predictors contribute independent variance. I demonstrate how the modeling procedure can be used to directly contrast the importance of various cues in launching speech in a series of simple production experiments, and compare the model outcomes to results from corpora.

Using this modeling approach, I show that in an experimental paradigm where planning is made trivially easy but the timing of speech launching is left uncertain, the most important predictors reflect simply the timing of the prior utterance. Using these alone in a predictive model accounts for most of the variance in turn gap, and reveals distributions of turn gap that look highly similar to gap estimates from corpora. This suggests that individuals attend to the prior utterance's timing more than any other cue in order to launch coordinated speech, and highlights a new way of reconciling results from various methodologies.

References:

- Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys*, 4, 40-79.
- Corps, R. E., Crossley, A., Gambi, C., & Pickering, M. J. (2018). Early preparation during turn-taking: Listeners use content predictions to determine what to say but not when to say it. *Cognition*, 175, 77-95.
- Magyari, L., & de Ruiter, J. P. (2012). Prediction of turn-ends based on anticipation of upcoming words. *Frontiers in Psychology*, 3, 376.
- Levinson, S. C., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology*, 6, 731.
- Zhang, Y., & Yang, Y. (2015). Cross-validation for selecting a model selection procedure. *Journal of Econometrics*, 187(1), 95-112.

The role of syntactic and prosodic cues in the anticipation of conversational turns

J.P. de Ruiter

Depts. of Computer Science and Psychology
Tufts University

Since the interest in conversational turn-taking took off in the 1970s there have been two fundamentally different approaches to the cognitive mechanisms responsible for the accurate temporal alignment of successive turns by different speakers. The first is what I will call the “signal approach”, pioneered by Starkey Duncan Jr. and Donald Fiske. The fundamental assumption of this approach is that the speaker produces turn-taking-relevant signals that the listener can exploit by responding appropriately. The alternative approach, which I will call the “projection approach”, suggested by the founding parents of Conversation Analysis, assumes that listeners are able to exploit the grammatical predictability in the speaker’s turn to anticipate the end of the current speaker’s turn before it occurs. The schism between these two approaches continues to influence the field to the present day, and obstructs progress. I will discuss these two approaches and some important theoretical and empirical arguments for and against them. Finally, I will suggest an approach that integrates and accommodates the different and seemingly contradictory findings from both approaches.

Social and cognitive explanations of dialogic resonance in everyday conversation

Nele Pöldvere, Victoria Johansson & Carita Paradis

Dialogic resonance, or the tendency for speakers to reproduce constructions across speaker turns, is a compelling type of coordination in everyday conversation. In (1), speaker B resonates with A along words, structures and meanings, while at the same time disagreeing with him. Moreover, as indicated by the square brackets, the disagreement is produced in overlap with A's prior turn, a turn-taking pattern that has previously been more commonly associated with agreement (Pomerantz, 1984).

- (1) A: she hasn't hitherto been particularly interested in religious things [has she]
B: [you mean] she hasn't particularly been up at seven AM

In this study, we take a closer look at *why* and *when* speakers make use of each other's contributions in dialogue. What are the social and cognitive aspects underpinning this behaviour? Previous work in two different areas of dialogue research offers different explanations. Du Bois (2014), on the one hand, argues that resonance is a socially motivated phenomenon that occurs for various communicative purposes, while Garrod and Pickering (2004) regard dialogue in general to be an automatic cognitive process facilitated by the reuse of prior expression. In this study, we aim to provide empirical evidence in support of the close association between social motivation and cognitive facilitation in dialogic resonance in everyday conversation.

We explore (i) social motivation through the functions that resonance has in discourse (agreement vs. disagreement) and (ii) cognitive facilitation through measurements of turn transitions. Firstly, based on previous research on the potential for resonance to occur in disagreement (Dori-Hacohen, 2017), we predict that

- resonance is more likely to be used by speakers in disagreement, while non-resonance is the preferred option in agreement.

Secondly, based on the view that the reuse of prior expression has a facilitating effect (Garrod & Pickering, 2004), we predict that

- transitions between speaker turns are faster in resonating sequences compared to when the sequences are constructed anew.

In addition, we explore whether or not there are timing differences between agreement and disagreement.

The sample contains 100,000 words of everyday face-to-face conversation among dyads from London–Lund Corpus 2 of spoken English. Using ELAN, we extracted 260 resonating and 316 non-resonating sequences, and analysed them in terms of function (agreement vs. disagreement) and the duration of turn transitions. Then, we fitted mixed-effects regression analyses to the data to test the predictions above.

The results provide full support for the first prediction. We propose that this is due to the cooperative nature of everyday conversation and the important role that resonance plays in mitigating the force of the ensuing disagreement through linguistic parallels. The second prediction also receives support. We see this as an indication that resonance has a facilitating effect on turn uptake, prompted by the activation of the same linguistic representations in the prior turn. Even though disagreements were expressed later than agreements, their mean durations were still strikingly fast (200–

300 ms). This leads us to conclude that, in resonance, the face-saving intersubjective motivations of resonance combine with its facilitating cognitive effect to promote smooth communication.

References

- Dori-Hacohen, G. (2017). Creative resonance and misalignment stance: Achieving distance in one Hebrew interaction. *Functions of Language*, 24(1), 16–40.
- Du Bois, J. W. (2014). Towards a dialogic syntax. *Cognitive Linguistics*, 25(3), 359–410.
- Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *TRENDS in Cognitive Sciences*, 8(1), 8–11.
- Pomerantz, A. (1984). Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In J. M. Atkinson & J. Heritage (Eds), *Structures of social action: Studies in Conversation Analysis* (pp. 57–101). Cambridge University Press.

Turn-taking in conversation from the larynx down

Marcin Włodarczak, Mattias Heldner

Department of Linguistics, Stockholm University
{wlodarczak, heldner}@ling.su.se

In this talk, we will give an overview of some of our results, both old and new, about respiratory and phonatory turn-taking cues. Both of these aspects of turn coordination are rarely addressed in literature, which focuses primarily on its articulatory and prosodic characteristics.

In the respiratory part of the presentation, we will discuss a new categorisation of turn-taking events which combines the criterion of speaker change with whether the original speaker inhales before producing the next talkspurt. We will demonstrate that the latter criterion could be potentially used as a proxy for pragmatic completeness of the previous utterance (and, by extension, of the interruptive character of the incoming speech). Specifically, respiratory properties of silences accompanied by speaker change in which the original speaker continues talking without breathing in are similar to those in within-speaker, turn-holding silences. We will also present evidence that the likelihood of speaker change is higher during pauses accompanied by a respiratory hold, suggesting that breath holds are used in reaction to incoming talk rather than as a turn-holding cue. In addition to analysing dimensions which are routinely omitted in studies of interactional functions of breathing (exhalations, presence of overlapping speech, breath holds), we will analyse patterns of breath holds in silent breathing and show that breath holds are sometimes produced towards the beginning (and towards the top) of silent exhalations, potentially indicating an abandoned intention to take the turn. We claim that the breathing signal can thus be successfully used for uncovering *hidden* turn-taking events, which are otherwise obscured by silence-based representations of interaction.

Moving up from the lungs to the larynx, in the second part of the talk we will focus on our ongoing work on voice quality variation in spontaneous interactions, a topic which has received little attention so far, not least because of the technical difficulties associated with recording phonation in continuous speech. In order to circumvent these problems, we are using miniature accelerometers attached to the skin of the tracheal wall below the glottis (“throat microphones”). The method, which has been used for some time in ambulatory postoperative voice monitoring, provides a good approximation of the voice source without the need for glottal inverse-filtering. We will demonstrate that the accelerometer signal can be successfully used to differentiate between voice qualities in isolated vowels while being unaffected by vocal tract resonances, f_0 and speaker variation. We will also present some preliminary results comparing several voice quality measures in speech intervals preceding silences accompanied by speaker change or followed by more speech from the same person. We demonstrate that utterances ending in speaker changes are characterised by lower periodicity and higher rates of creaky voice. The findings are thus consistent with the “trailing-off” character of these silences, as suggested in literature.

When two are “doing being” one: Co-animation and association in collaborative turn sequences

Marina N. Cantarutti - The Open University (UK) - marina.cantarutti@open.ac.uk

In their description of the systematics of turn-taking, Sacks, Schegloff, and Jefferson (1974) discuss how in the highly-coordinated and organised turn-transitions between participants there is a minimisation of gap and overlap, and how next speakers are allocated. Important to this description, moreover, is also that these participation “rules” are actually not centred around the *number of participants*, but rather, around the *number and organisation of parties* (Schegloff, 1995). This entails an independence between participation roles and their incumbency (Levinson, 1988), and thus, the reality that conversational parties may temporarily be inhabited by more than one interlocutor (Bolden, 2013; Schegloff, 1995), a process called “association” (Lerner, 1993). When participants associate, they explore the possibilities of the turn-taking system to co-construct an interactional activity as one party. They can do so as self-selected representatives of a collectivity (Schegloff, 1995), while teaming up for Other-repair (Bolden, 2013; Ekberg, 2012), or in collaborative turn sequences and turn-sharing (Lerner, 2002, 2004), where a participant completes or continues a co-participant’s turn simultaneously or adjacently.

This paper scrutinises co-participant co-constructions of jointly-owned sequences through another collaborative practice that further blurs the boundaries between Self and Other: *co-animation* (Cantarutti, 2018, 2020), that is, the joint quoting or (re)enactment of the same voice addressing an absent third party (see figure 1). A close inspection of these finely-coordinated practices reveals that first animators build syntactic, semantic, and prosodic projectability into their turns, framed with intricate gestural and gaze patterns that co-participants are seen to visibly orient to and use as “opportunity spaces” (Lerner, 2004) to co-animate. Co-animators complete or continue prior animations with no quotative preface and build coherence based on features of the first animation while overlaying aspects of their independent design. They co-construct a voice through prosodic and gestural integration, continuing prior stretches of perceptually isochronous rhythm and beat gestures, or matching changes in articulatory settings and postural-gestural configurations.

This paper provides evidence of how these multimodal forms of mutual orientation, coupled with the sharing of an animated voice, contribute to a complex view of turn-transition and turn-sharing that a) recognises that responsive turns by B may be timed and designed as retrospectively belonging to prior turns by A; b) sees the use of conversational resources as distributed, and c) considers other levels of lamination that put turn-taking at the service of relational concerns like association.

This study is based on 40 cases of co-animation in English interaction, and adopts a multimodal perspective to the empirical, inductive, and microanalytic methods of conversation analysis and interactional linguistics (Couper-Kuhlen & Selting, 2001), combined with an impressionistic plus instrumental study of phonetic detail (Local & Walker, 2005) and the annotation of gesture based on phases and gesture types (McNeill, 1992, Kendon, 2005).

Figure 1. MCY20MUG “Scoon” (GAT2 Fine, adapted; Mondada Multimodal Transcription; co-animation in bold)

```

02 SAM you should tOtally go to ^Edinburgh ^CASTle, (or)
03     (or into) ^STIRling?
04     i really `LIKE stIrling [^CASTle?]
05 EVI                                     [ x xx   ] to Edinburgh ^CASTle?
06 SAM theyre `BOTH like; = com`PLETE; (.) CASTles?
07 EVI yEah `NO. = i wEnt to `Edinburgh.
08     but i dIdnt (Enter) the ^CASTle?
09 SAM ((lip smack)) `YEAH. = theyve gOt the:ir stOne of ^DESTiny?
10     (0.7) ((click)) (1.2)
11     the `s:tOne of `SCOON; h°
12     its spElled ^SCONE, = but ↓yOu say ^SCOON,
13 EVI <<cr, l, p>`SCOON is bEtter than ↓`SCO(?h°)NE.>[hh°      ]
14 SAM                                     [^YE:A:H,]
15 EVI Oh i `LOVE i?; ((click))
16EVI *Is it +*`SC*ONE*+ ((skɒn))           = *or           +↑`SCONE*+ ((skəʊn)).
     evi *RH point-,,,,,*                               *raises RH>
           *turns and looks R                               *turns L.-----*turns R>
     sam           +rubs RThumb+raises RThumb           + rubsRT +raises RT
17     `NO.+ * (.) [<<f>*`S::COON ((ske:n)) >]*2h°aha
     evi >*RH point,,,,,,,*
           >*turnsR *leans forward-----* 
     sam           +raises RT
18 SAM           [<<f>+`S::COON ((ske:n)) >]+
     sam           +raises folded RT----- +
19     [yEah ^YE:AH? haha]
20 EVI [hahahahah          ]

```

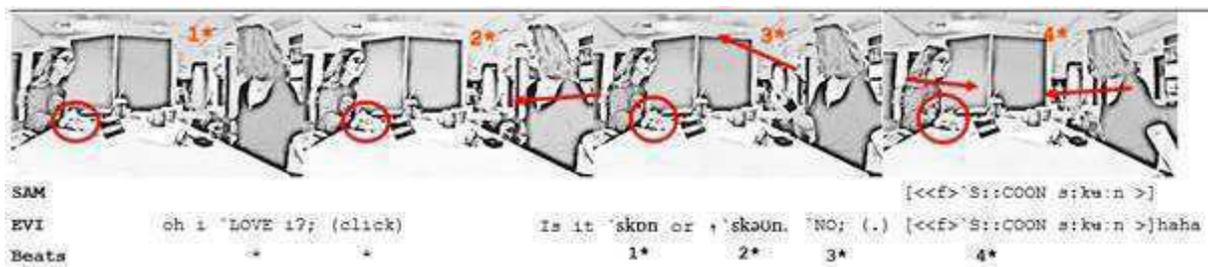


Figure 2. Multimodal partiture with beat gestures by Evi and orientation by Sam (lines 15-18)

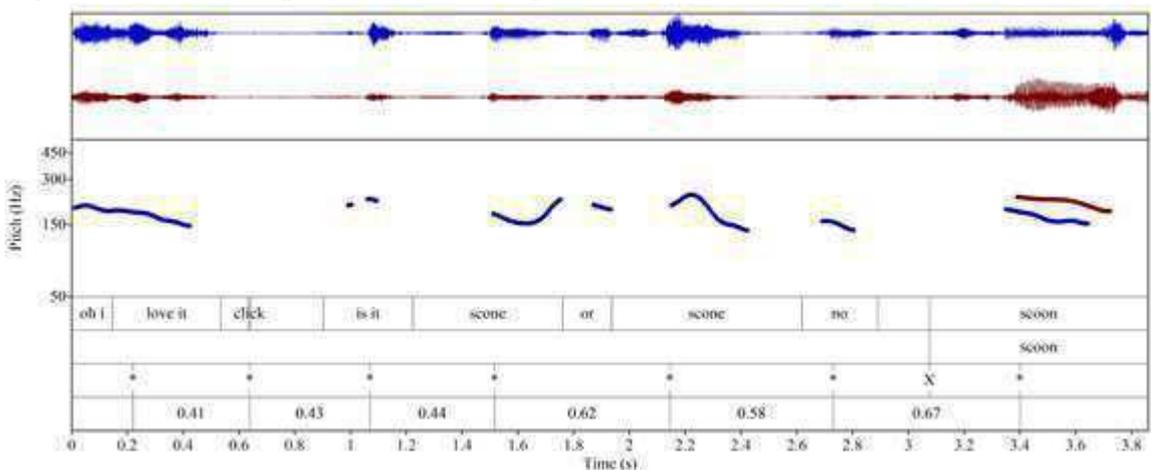


Figure 3. Waveform, f0 trace (in Hz logarithmic), and rhythmic intervals of lines 15-18)

Turn-final lengthening is really rallentando across (much of) the turn

Christoph Rühlemann & Stephan Gries

Turn transition in talk-in-interaction is achieved with remarkable precision, most commonly following a gap of no more than 200 ms (e.g., Stivers et al., 2009). How the precision is achieved is a complex issue given the wide range of variables conversationalists deploy to project (as speakers) and predict (as listeners) turn completion. This paper aims to contribute to a deeper understanding of one such variable used by speakers to project turn-completion: changes in word duration in turns-at-talk. As word duration varies significantly due to influences from a large number of confounds, we approach the challenges inherent in “[p]roviding robust, quantified, comparative measures of duration” (Local & Walker, 2012: 259) by fitting mixed-effects models based on naturally occurring corpus data from the conversational subcorpus of the British National Corpus. Contrary to previous research, which hailed the turn-final drawl as a turn-yielding cue, the models indicate that drawling, or rallentando, affects not just the turn-final syllable/word but large portions of the turn depending on the placement of nuclear stress. Our models suggest that rallentando constitutes, not a one-off cue marking the turn’s end-point upon its occurrence, but an extended process advance-projecting the turn’s durational envelope. Also, as a graded advance-projecting resource, rallentando is in and of itself insufficient to signal turn completion reliably; listeners are likely to rely on turn rallentando in unison with other, preferably discrete cues marking the turn-completion point upon its occurrence, for “recogniz[ing] that a turn is definitely coming to an end” (Levinson & Torreira, 2015: 12) and triggering the launch of the next turn.

Local, J., & Walker, G. (2012). How phonetic features project more talk. *Journal of the International Phonetic Association*, 42, 255–280.

Levinson, S. C., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology*, 6, 731. <https://doi.org/10.3389/fpsyg.2015.00731>.

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., de Ruiter, J. P., Yoon, K.-E., & Levinson, C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of the Sciences*. U.S.A., 106(26), 10587–10592. <https://doi.org/10.1073/pnas.0903616106>.

Horváth Viktória – Krepesz Valéria – Gyarmathy Dorottya – Huszár Anna – Hámori Ágnes: Some new phonetic aspects of smooth turn taking in triadic conversations in Hungarian

The organization of turns and turn-taking is a central phenomenon of conversations (Sacks et al. 1974, Levinson–Torreira 2015). One of the main questions is, how can it be so smooth? Furthermore, which mechanisms and cues take part in managing turn-takings if there are more than two participants in the conversation? This paper presents data on phonetic parameters of turn takings and the speech of the current speaker related to the support of smooth turn transitions, based on the results of corpus-based analysis on Hungarian triadic conversations. The research analyses the phonetic context and timing features of the current speaker's utterances that show differences between the appearance in a position near to the turn taking and in a position farther from the turn-takings.

The material of the research consists of 20 conversations from the sub-corpus of the Hungarian Spontaneous Speech Database (cf. Horváth et al., 2019), containing informal three-party conversations annotated in Praat (where two of the three participants were the same and only one speaker changed during the recordings). We explored a) the frequency, the type (turn taking after pause or after overlapping speech) and the timing patterns of turn-takings (FTO-values – Floor Transfer Offset, cf. Ruiter et al. 2006), b) dynamic changes of phonetic properties during the conversation related to the distance to the upcoming turn end, c) differences of the examined phonetic factors in the positions near to the turn-takings and further from the turn-takings. In the case of intervals' position, the centre point of the interval was taken into account (all in the case of silent pauses, backchannel responses, and articulation tempo). To examine the dynamic changes, we assigned interval-specific average values to the time of the centre of the interval and these were compared with each other on the different point of the conversation. The distance of the centre of the intervals from the pre- and post-turn-takings were examined, and then classified into 2 clusters (intervals nearer or further from TTs). The articulation rate in each interpausal unit (altogether appr. 3000 items), the frequency and duration of the silent pauses were measured in using Praat script as well according to their position in the conversations. In addition, backchannel responses of other interlocutors were also analysed with regard to their role in cooperative management of speaker change. Mixed models, cluster analysis and correlation analysis were carried out for the statistical analysis in R program.

Results showed the majority of turn-takings occurred with positive FTO-values with a similar range to other languages (cf. Stivers et al. 2009). Turn-takings realized more likely with silent pauses at the beginning of conversations, while turn-takings realized with overlaps in a greater ratio at the end of conversations. Differences were found in speech parameters in the vicinity of turn-takings: the duration of silent pauses were shortened, the value of articulation tempo decreased. Furthermore, backchannel responses were more frequent before turn-takings than following them, confirming the cooperation of the conversation partners in managing turn-taking. Our results corroborate that the turn-taking system is jointly coordinated by the participants: the dynamic changes of speech parameters of the current speaker as well as the backchannel responses of the listener are important cues of managing turn-takings smoothly. The results support that the current speaker uses multiple acoustical cues to indicate the boundary of the utterances. Furthermore, results provide new information about the dynamic changes of phonetic parameters in the conversation.

References

- De Ruiter, J. P., Mitterer, H., & Enfield, N. J. (2006). Projecting the end of a speaker's turn: a cognitive cornerstone of conversation. *Language*, *82*(3). 515–535.
- Horváth, V., Krepsz, V., Gyarmathy, D., Hámori, Á., Bóna, J., Dér, Cs. I., & Weidl, Zs. (2019). Háromfős társalgások annotálása a BEA-adatbázisban: elvek és kihívások [The annotation of three-participated conversations in Hungarian Spontaneous Speech Database: principles and challenges]. *Nyelvtudományi Közlemények*, *115*(1). 255–274.
- Levinson, S. C., & Francisco, T. (2015). Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology*, *6*.
- Sacks, H., Schegloff, E., & Jefferson, G. (1974). A simplest systematic for the organization of turn-taking for conversation. *Language*, *50*, 696–735.
- Stivers, T. et al. (2009). *Universals and cultural variation in turn-taking in conversation*. Proceedings of the National Academy of Sciences of the United States of America, *106*(26). 10587–92.