Beyond Philology No. 19/3, 2022
ISSN 1732-1220, eISSN 2451-1498

https://doi.org/10.26881/bp.2022.3.03

The facilitatory role of utterance splitting in text-based CMC

MAREK PLACIŃSKI

Received 15.11.2021,
received in revised form 13.11.2022,
accepted 16.11.2022.

Abstract

This paper focuses on the role of the structure of turns in quasi-synchronous text-based computer mediated-conversation. Prior research has found that interactants in this type of communication submit their messages in two ways: either as a long single message or as a sequence of shorter postings. We hypothesize that the latter strategy – called utterance splitting – facilitates communicating online both in terms of turn-taking (initiation of repairs and holding the floor) and sentence processing (predicting informational content and lowering entropy). To evaluate the hypothesis, an experiment in which naive participants interacted with a confederate writer was conducted. We found that although communicative success did not depend on the condition to which the participants were assigned, the conversations in which the confederate sent split utterances were, on average, shorter and required less words to communicate the same intention.

Keywords

computer-mediated communication, informational entropy, turn-taking, sentence processing

Rola strategii utrzymywania tury konwersacyjnej w quasi-synchronicznej tekstowej komunikacji internetowej

Abstrakt

Niniejszy artykuł rozważa wpływ formy tur konwersacyjnych w quasi-synchronicznej tekstowej komunikacji internetowej. Wcześniejsze badania zauważyły, że rozmówcy w owym rodzaju komunikacji realizują tury konwersacyjne na dwa sposoby: bądź w postaci jednej długiej wiadomości, bądź w serii kilku krótszych wiadomości. W tym artykule stawiana jest hipoteza, że druga strategia ułatwia zarządzanie procesem konwersacji (przewidywanie końca tury, inicjowanie napraw konwersacyjnych) i obniża koszty związane z przetwarzaniem zdań (poprzez stopniową integrację wiadomości oraz zmniejszenie entropii informacyjnej). Aby zweryfikować hipotezy, przeprowadzono jedno badanie. O ile eksperyment nie wykazał istotnego statystycznie wpływu formatu tur konwersacyjnych na sukces komunikacyjny, o tyle dane sugerują, iż format ma istotny wpływ na wspólny wysiłek komunikacyjny – na korzyść krótszych wiadomości.

Słowa kluczowe

komunikacja internetowa, entropia informacyjna, przejmowanie tur konwersacyjnych, przetwarzanie zdań

1. Introduction

Text-based computer-mediated communication (henceforth CMC) is one of the most popular means of interaction online. However, it differs from face-to-face interaction to a large degree in terms of turn management. This paper explores the notion of utterance breaks, i.e. the division of a message into smaller units (see section 4) and argues that utterance splitting is a mechanism that facilitates text-based communication.

2. Synchronous text-based CMC

Written CMC occurs via standalone clients, such as WhatsApp or various IRC channels, or is part of other services, for example, Facebook or Google. However, due to technical constraints, it differs significantly from face-to-face talk (Garcia & Jacobs 1999, Herring 1999) both in terms of synchrony and multi-modality. With respect to synchrony, Baron (2010) suggested the typology presented in Table 1.

Table 1
Baron's (2010) typology of CMC

one-to-one	synchronous	asynchronous
	instant messaging	emails, texting
many-to-many	chat, computer conferencing	bulletin boards, listservs, blog sites

Hence, CMC can be classified based upon the expectations associated with response time. Certain types are called "synchronous" because there is an expectation of an immediate response – such as in instant messaging services. On the other hand, if no such expectation is held, the service is classified as "asynchronous". A good example of asynchronous CMC is email – generally, if one sends an email message, they do not expect to receive an answer immediately, quite to the contrary. Still, it does not mean that the classification is rigid, and especially in text-based chats the boundaries are fuzzy: a reply to an email may arrive virtually immediately, whereas a response for a message in an instant messaging chat can take some time. The same logic applies to the number of participants involved.

However, Baron's (2010) classification did not take into consideration another important factor in classifying CMC, namely by what means interactants communicate. In addition to the previous typology, we can also classify CMC on the basis of the

type and number of channels, and their modalities[1], involved to communicate a message. From this perspective, CMC can be divided into types that engage primarily the visual, primarily the audio, or both sensory modalities (see Table 2).

Table 2
Mediated communication classified according to the modality involved

	monomodal	multimodal
primarily visual	instant messaging, group chats, emails, texting	video conferencing, video calls
primarily auditory	phone calls, recorded messages	

The new classification should not be viewed as one that rejects Baron's (2010), but rather an addition that had previously been overlooked. Here, we develop an argument that in addition to the degree of synchronicity, the primary modality also influences the structure of turns and the process of turn-taking. We build on Clark's (1996) principle of least collaborative effort to explain interlocutors' strategies of turn-taking in text-based CMC, by which they trade off the costs of communicating in a highly constrained environment.

3. The synchronicity of CMC revisited

The degree of synchronicity of text-based CMC has been a subject of debate. Interacting via instant messaging applications is never *as synchronous as* face-to-face conversation. Even in the world of mediated communication, text-based communication lags behind telephone conversations and video conferences in terms of synchronicity. Texting, instant messaging and group chats involve two or more people that engage in a conversation,

[1] Modality herein is defined as the sensory modality used in perception (see Zlatev et al. 2017).

but without these individuals being co-present: they send and receive messages from different places in the world. Hence, should instant messaging actually be labelled *synchronous*?

This difference between various kinds of mediated communication led certain researchers to use the label *quasi-synchronous computer-mediated communication* (henceforth QS-CMC) to refer to instant messaging and text-based group chats (Garcia and Jacobs 1998, 1999, Zemel 2005). Garcia and Jacobs motivated adopting the term by stating that "although posted messages in [QS-CMC] are available synchronously to participants, the message production process is available only to the person composing the message" (1999: 339). That is to say, in most text-based CMC, the message can be read and processed only once it has been produced, edited and submitted by the current writer. It is up to the current writer to decide whether the message is complete and can be submitted to the chat. This fact stands in stark contrast with face-to-face interaction, where the listener witnesses the production of a message *on-line* – as it is being uttered by the current speaker.

4. Interaction management in QS-CMC

Despite the fact that both instant messaging and face-to-face conversation are instances of interpersonal communication that involves at least two participants, they differ in terms of resources available for interaction management. Whereas in face-to-face talk, interlocutors can rely on sentence structure (Sacks et al. 1974), intonation (Beattie et al. 1982, Cutler and Pearson 1983, Oliveira and Freitas 2008, Schaffer 1983), tempo (Rühlemann and Gries 2020), gestures and gaze (Bavelas et al. 2002, Gambi et al. 2015, Kendon 1967, Rossano 2012, Żywiczyński et al. 2014), as well as a combination of these features (Ford and Thompson 1996), these signals are limited in QS-CMC due to the lack of co-presence. Together with the data handling protocol used in CMC (FIFO: first-in, first-out), they lead to the lack

of simultaneous feedback and disrupted turn adjacency (Herring 1999).

Both of these factors have an impact on turn-taking in text-based chats. Overlaps, existent but infrequent in face-to-face communication, are completely out of the question in QS-CMC (Garcia and Jacobs 1998, Herring 1999). This is so due to the second technological factor mentioned above – the fact that the server protocol manages messages in a sequence one at a time. Therefore, new messages always appear on the screen in an orderly fashion. However, what does occur in QS-CMC is an overlap of exchanges, especially in multi-party types of text-based CMC. There, different conversational threads are intertwined, meaning that messages from conversation A may interrupt messages from conversation B by appearing between its adjacency pairs (Garcia & Jacobs 1998, Gibson 2014, Herring 1999, Simpson 2005). With regards to gaps (or pauses) that do not occur often in face-to-face communication, they do appear in QS-CMC, either because of the time it takes the addressee to read a message and compose his or her own reply, or because of lags (temporary delays) of the server (Anderson et al. 2010, Herring 1999, Riordan et al. 2013).

Furthermore, some (Gibson 2014, Herring 1999, Schönfeldt and Golato 2003) have observed that turn-alternation is not as smooth as in face-to-face conversation. Whereas in such interactions the first pair-part usually invites one second pair-part that is temporally adjacent, QS-CMC conversations can involve multiple responses to one utterance that are not necessarily spatially adjacent. This often results in disruption of adjacency pairs, where two turns are spatially adjacent (in a sequence, one after another), but are functionally unrelated – they do not belong to the same pair. Garcia and Jacobs (1999) call this feature of text-based chat *phantom adjacency pairs*. Although apparently problematic, these displaced utterances enable interlocutors to do something that is not possible in face-to-face conversation: directly reply to any previous utterance. Phantom

adjacency can also occur as a result of many conversational threads intertwining with one another.

Consider a typical transcript of a multi-participant chat – an Internet Relay Chat (IRC) conversation) – obtained from a support channel for users of the Linux Ubuntu operating system (Uthus and Aha 2013). The channel includes experienced users that can offer help to solve a technical problem, and those users that seek help.

> [01:19] <designbybeck__> oh i was looking at c instead of F.... yeah these are running hot and I'm only on chat!!
> [01:19] <designbybeck__> i closed Firefox with Facebook open
> [01:19] <designbybeck__> so now just IRC
> [01:19] <designbybeck__> and the fan is running and burning hot to the touch!!!!
> [01:20] <WeThePeople> how do i get wireless working in fluxbox
> [01:20] <puppy_parade> for a laptop, shooting the fan with canned air usually does wonders.
> [01:21] <designbybeck__> thanks puppy
> [01:21] <puppy_parade> or blowing it it to get dust out almost as well
> [01:21] <puppy_parade> (with the machine off)
> (Uthus and Aha 2013)

This fragment of a conversation on the channel involves several interlocutors talking in different conversational threads – despite being posted to the same channel, they are not related and may involve different interactants. In this fragment, <designbybeck_> runs into problems with laptop overheating while running Linux Ubuntu. The user is being aided by <puppy_parade>, who predicts that the temperature problem stems from hardware problems (perhaps too much dust) and suggests clearing the fan with compressed air. Their adjacency pairs are interrupted by <WeThePeople>, who asks a question unrelated to the conversation between <designbybeck_> and <puppy_parade>.

The absence of the non-verbal features mentioned above, including the lack of possibility of online language processing, means that interlocutors in QS-CMC must adapt to the

technology via which they communicate. One of such strategies might be utterance splitting, a technique developed to overcome obstacles of interaction via QS-CMC. The phenomenon of posting two messages to the chat has been observed independently by different researchers, in effect receiving different labels. For instance, Baron (2010) called it *utterance splitting*, Anderson et al. (2010) *delayed completions*, whereas Tudini (2013) *turn incrementation*. Here, to avoid any further confusion, we will use the term *utterance splitting*. Several split utterances submitted by the same user and continuing the same theme, and generally not interrupted by other interlocutors, are henceforth called *sequences*. As such, attempts were also made to explain the function of splitting utterances in QS-CMC.

4.1. Split utterances as intonation units

Baron (2010) analysed utterance splitting from the perspective of Chafe's (1980, 1987) intonation units. According to Chafe (1987), dialogue is composed of concepts – small units that correspond to "new" and "old" ideas. In his view, the carrier of these concepts and the basic unit of discourse in speech are intonation units, preceded and followed by a pause and uttered under a single intonational contour. Chafe (1987) argued that those pauses between intonational units signify cognitive processes underlying the shift from one concept to another.

Having noticed utterance splitting, Baron (2010) drew a comparison between such a means of organising utterances in written communication and intonation units in speech, as both of them serve to divide a larger stretch of discourse into smaller pieces. However, as she noted, there are differences between utterance splitting and intonation units: an intonation unit can be signalled by a pause, a clause-final intonation contour, usually begins with a conjunction, or syntactically is a single clause, the only means of chunking an utterance in CMC is sending an entire message in several sequences. Therefore, although not ideally corresponding to one another, intonation

units nevertheless serve as a good point of departure for the analysis of utterance splitting.

Baron (2010) analysed dyadic conversations from IM chat, compiled in 2003 and involving 22 college students. The sequences taken into consideration involved those whose first message was, for instance, a subordinate clause and the other the main clause. Consider the following example:

> [01:19] <designbybeck__> i closed Firefox with Facebook open
> [01:19] <designbybeck__> so now just IRC
> [01:19] <designbybeck__> and the fan is running and burning hot to the touch!!!!

The extract consists of three messages submitted by the same user, uninterrupted by others. The sequence, as a whole, is a compound sentence: the first message is formally a clause, the first conjunct of the coordination. The second message is an adjunct. The third is the second conjunct that begins with the coordinating conjunction *and*.

In total, Baron (2010) found 454 sequences with multiple messages, 132 contained utterance breaks (nearly 30 %). In addition, conversations in male and female dyads also revealed that men are more likely to use such sequences than women ($p < 0.0001$). With respect to their structure, sequences were predominantly composed of complex and compound sentences, independent clauses, although some of them also included adjectival, adverbial, noun and verb phrases.

The analysis led Baron (2010) to two conclusions. Primarily, because of the relative brevity of turns in the corpus, common use of single-word messages and long closings, instant messaging resembles spoken communication (rather than written). Second, the study revealed an incompatibility between Chafe's (1980, 1987) intonation units and utterance splitting in instant messaging. Although there were certain common features, like the brevity of messages in sequences and in utterance units and pauses occurring before conjunction, there were other striking differences. Utterance breaks did not occur frequently before

noun or verb phrases in Baron's (2010) corpus, whereas pauses in Chafe's (1980) study do appear before them. Another difference results from the type of activities spoken dialogue and written interactions are: whereas dialogue is foregrounded, i.e. it is the main focus of both parties involved, it is not the case for written conversation. Frequently, Internet users are engaged in multiple other online activities on top of conversation. Thus, whereas pauses are relatively short and gaps hardly ever occur, the time between two adjacency pairs or two elements of a sequence can span from short pauses to gaps as long as 5 minutes (Baron 2010).

4.2. Split utterances as delayed completions

The turn-taking mechanism in CMC was also investigated by Anderson et al. (2010). There are two caveats of this study, however. Primarily, although the study is relatively recent – it was published in 2010 – the corpus for the study was gathered in the 1990s. At the time, relatively few users had access to the Internet and participants of the experiment may have developed communication strategies 'on the fly'. Secondly, the participants communicated over a two-way protocol, which, unlike most popular messaging services, displays the message as it is being produced by the current writer.

The study involved an analysis of one videotaped text-based conversation. Recording the participants made it possible to detect typing periods, pauses, and overlaps, which normally are not studied in this fashion in CMC research, where overlaps technically are out of the question (as two messages cannot be posted at exactly the same time) and pauses are measured as periods between submitting two messages.

Overlaps, unlike in face-to-face communication, occurred very frequently. This means that two writers could start typing at the same time (self-selection) or one writer started to type at a time when the other has not finished yet. According to Anderson et al. (2010), such an overlap is mitigated by *delayed*

*completion*s, i.e. by a strategy in which the current writer divides a message into two portions, waiting with typing its second part until the other participants have read the message and are prepared to read more. Hence, the readers may attach a portion of an utterance to previous messages in the chat history, therefore reconstructing some holistic meaning from individual parts. Attaching one utterance to another is greatly facilitated by the fact that there is no *rapid fading* in text-based QS-CMC: a user can always scroll back to a previous utterance, in contrast with face-to-face communication, where sounds of speech disappear once they are uttered.

Concerning pauses, users were observed to apply them systematically in order to manage interaction. As Anderson et al. (2010) observed, "[t]he participants appear to be engaging in intermittent talk followed by strategic pauses throughout the data corpus". This means that, unlike in face-to-face interaction, pauses occur often, and they perform a specific function in chat: pauses trade off the communicative costs incurred by the frequently appearing overlaps. Pausing in two-way communication systems also means that others have ample time to read whatever the writer has typed in the chat box.

All in all, we can see that Anderson et al. (2010) observed pauses in typing messages, which are similar to Baron's (2010) split utterances. Anderson et al. (2010) infer that splitting utterances in this fashion is an adaptive strategy that gives the addressee(s) a chance to process the utterance before the writer continues with the second part of the message. Thus, the problem of overlapping utterances is at least partially mitigated, and utterance splitting may explain why users manage to communicate successfully despite the incoherence of QS-CMC.

4.3. Split utterances as turn extensions

As mentioned previously, Baron (2010) observed that intonation units split a spoken utterance into smaller parts. However, in face-to-face interaction, a speaker can also convey a message

over several complete turns that are not interrupted by other speakers by means of *turn extension*. Such turn extensions occur at transition relevance places where a speaker elects to continue speaking (self-selects), and they typically elaborate on what the current speaker has said previously. According to Tudini (2015), turn extensions are used in the same way by interlocutors in QS-CMC – utterance splitting in QS-CMC is a means to build on the meaning of previous turns.

Tudini (2015) explores turn extensions from Schegloff's idea (2000 in Tudini, 2015) idea of symbiotic relation between the first and subsequent messages of a sequence and focuses on their syntactic relationship. The material in the study consisted of 12 conversations between native and non-native speakers of Italian. The speakers communicated in their free time, as an out-of-class activity that was supposed to improve their language skills. Such a study design ensured ecological validity of the data, since participants interacted in 'regular' circumstances.

Splitting utterances, according to Tudini (2015), might be an attempt to hold the floor in conversation and deny the interlocutor the possibility to post. Extensions, because they can be written faster than whole messages, can also be used in order to maintain the adjacency of turns and decrease the probability that some other speaker will intervene. Splitting turns may also be a sign that someone "is politely attending to the conversation" when the parties are not engaged in the conversation at the same time (Tudini 2015: 649). The last reason she enumerates is the fact that certain groups display a preference for shorter turns; therefore, an individual may align with the preferences of the group. Apart from enumerating reasons for splitting an utterance, Tudini (2015) also describes their functional aspect in conversation.

4.4. A comparison

The papers by Baron (2010), Anderson et al. (2010) and Tudini (2015) analysed the same phenomenon, albeit from different perspectives. All three analyses focused on the problem of turn structure in QS-CMC and concerned themselves with the fact that interlocutors tend to split their utterances over several postings. Baron (2010) attempted to compare utterance splitting with intonation units in spoken dialogue, taking as a departure point Chafe's (1980) intonation units. The comparison failed, and the two phenomena were found to share only one feature – length, since both intonation units and split utterances are relatively short. However, splits tend to occur in different places within a message than pauses in spoken dialogue. As such, the second feature of intonation units – carrying new conceptual content – remains an open question.

The studies by Anderson and colleagues (2010) and Tudini (2015) are the closest to the purposes of this paper. Anderson et al. (2010) argued that the purpose of delaying the completion of an utterance is to provide other users with ample time to process the utterance. In such a way, splitting the utterance mirrors spoken interaction, where the addressee can hear an utterance as it is being produced, and not only its final version. Hence, utterance splitting can be said to facilitate (reduce the costs of) interacting via monomodal, quasi-synchronous text-based communication.

5. The cost of interaction: The principle of least collaborative effort

Anderson and colleagues (2010) argued that delayed completions may be used to facilitate communication. The works by Clark and Wilkes-Gibbs (1986) Clark (1996), Clark and Brennan (1991) provide a theoretical basis from which the issue of effort in conversation can be explored. Under this theory, conversation is a collaborative effort of the parties involved that pertains

to both coordinating the content of communication – what interlocutors are talking about – and the process – turn-taking and updating mutual knowledge (common ground) on a turn-by-turn basis. Common ground is updated by a process of grounding, i.e. building a body of knowledge based on what has been said earlier in conversation. An important feature of grounding is that the process is *medium-specific*: communicative context (such as interacting face-to-face or via a computer terminal) influences grounding.

Grounding is practically achieved by interlocutors making contributions in conversation: they propel it by exchanging turns. According to Clark and Brennan (1991), people tend to be very economical in making their contributions. The phenomenon was observed before Clark and Brennan's works, notably in Grice's work on the cooperative principle (Grice 1975). The two relevant maxims to effort in conversation are the maxim of quantity (contribute no more information than is necessary in conversation) and manner (be brief). Speakers, therefore, are supposed to produce *proper utterances*, i.e. utterances that will be easily understood by addressees. However, thus construed principle of collaborative effort assumes flawless contributions, which is not always the case (see the example below). According to Clark and Wilkes-Gibbs (1986), there are three problems that render the principle of least effort implausible: time pressure, errors and ignorance.

> Alan: Now -um do you and your husband have a j-car?
> Barbara: - have a car?
> Alan: Yeah.
> Barbara: No.

Were speakers always obeying the Gricean maxims, Alan would have made sure to deliver the message clearly, and the whole conversation would have featured much less repetition. Yet, this is not the case, and conversations are characterized by a high degree of repetition (Pickering and Garrod 2004). According to Clark and Wilkes-Gibbs (1986) what contributes to *improper*

utterances are features that are characteristic of interaction: time pressures, errors and ignorance. Since the principle of least effort is not compatible with naturalistic data, Clark and Wilkes-Gibbs (1986) formulated the principle of *least collaborative effort*:

> *The principle of least collaborative effort*: In conversation, the participants try to minimize their collaborative effort – the work that both do from the initiation of each contribution to its mutual acceptance.

The principle, according to Clark and Brennan (1991) explains a number of phenomena in conversation. For instance, the preference for repairs: speakers generally prefer to repair the utterance and initiate a repair on their own, rather than to rely on the interlocutor to repair or prompt the repair. Self-repair typically takes fewer turns in comparison with waiting for other interlocutors to initiate it. By the same token, speakers can produce their utterance in smaller 'chunks', anticipating acceptance after each small item.

Communication costs depend on the medium of communication. What is effortless in one medium can be more costly in another, or even completely out of the question (like pointing one's finger at an object in a letter). The absence of one of the features makes interlocutors rely on others. Clark and Brennan (1991) enumerate 8 features that affect communication: copresence, visibility, audibility, cotemporality (language production and perception occur at roughly the same time), simultaneity (the speaker sees the addressee's reaction in real time), sequentiality (turns appear in an orderly fashion), reviewability (the opposite of Hockett's (1958) rapid fading) and revisability (the speaker/writer can revise messages before sending them).

Clark and Brennan (1991) associated the features of cotemporality, sequentiality and reviewability with QS-CMC. The characterisation is a bit outdated, though, as these are qualities of two-way conferencing systems, not popular nowadays. Therefore, the list should also include revisability because, in one-

way communication protocols, writers can currently edit their messages before submitting them to the chat. Also, sequentiality is only weakly obeyed, especially in multi-participant chats, and cotemporality is not a necessary prerequisite for initiating a conversation.

Grounding costs can be divided into the ones that are incurred by the speaker (writer), by the addressee(s), or by all parties involved. The costs that are incurred by the speaker are formulation (formulating and re-formulating an utterance) and production costs (the actual act of speaking or writing). On the part of the addressee, there are reception (perceiving the linguistic signal) and understanding costs (understanding the content). Among the costs that can be associated with both the speaker/writer and the addressee(s) Clark and Brennan (1991) specify delay, asynchrony, speaker change and display costs. The first type of cost is related to the costs incurred by delaying an utterance to plan, revise and produce it more carefully. The second type are the costs associated with timing utterances incurred when parties communicate via non-synchronous media of communication, such as QS-CMC. Speaker change costs are the costs of turn-taking: whereas they are low in face-to-face communication, they are higher in QS-CMC because there are fewer cues that speakers can rely on in order to aid transition between speakers. Finally, display costs are connected with using gestures, facial expressions and pointing. Relatively effortless in face-to-face communication, these are completely out of the question in text-based QS-CMC.

6. Motivation

Anderson et al. (2010) hypothesized that utterance splitting may be a technique used by interlocutors to facilitate communication. The works by Clark and colleagues cited in the section above provide a theoretical background for Anderson and colleagues' (2010) conception: grounding, i.e. the process by which interlocutors update their mutual knowledge about the

situation under discussion, incurs costs. These costs vary according to the medium, and interlocutors involved in communication via particular mediums use resources available there accordingly to minimize the costs of reaching common ground, as the principle of least collaborative effort predicts.

Why would utterance splitting facilitate communication? This is due to the constraints on QS-CMC and an attempt to trade off the costs of interaction. Interlocutors in QS-CMC are not co-present and can neither hear nor see each other, which negatively impacts coordinating the content and process of communication, which is evident in transcripts of text-based CMC. As such, producing shorter utterances ensures that the other interlocutor does not take the floor prematurely. Another important feature of short utterances is their capacity for repair: if the addressee does not understand the writer's intention, they can signal it at an early stage. Therefore, a single part of a longer sequence of messages has to be reformulated. In an alternative extreme scenario, where the writer sends a whole 'paragraph' to the chat, repair would be much more difficult because the addressee would have to indicate the part which is difficult to understand, which in turn would lead to another coordination problem.

7. Experiment

To test the hypothesis that utterance splitting facilitates communication in QS-CMC, we adapted the Map Task experiment (Anderson et al. 1991). Twenty-two participants were enrolled to participate in the experiment via a university mailing subscription list. The participants interacted with a confederate writer through the Google Hangouts instant messaging app. All of the participants were native speakers of Polish and communicated with the confederate using this language. Due to Coronavirus restrictions on travel, participants did not take part in the experiment in the lab, but interacted from their homes. Each participant received instructions a day prior to the experiment.

7.1. Method

The Map Task is an interactive task designed to elicit dialogue (Anderson et al. 1991, Thompson et al. 1993). Originally, the experiment involved two participants who were instructed to communicate verbally in order to reproduce the path on the map. The participants were assigned the role of the leader (the participant with the route) or of the follower (the participant without the route). In the original experiment, each map had a similar layout, though particular landmarks differed – they were printed in different places or altogether absent on one map. The maps were translated into Polish for the purposes of this experiment. Figure 1 shows the two versions of maps used in the current experiment.

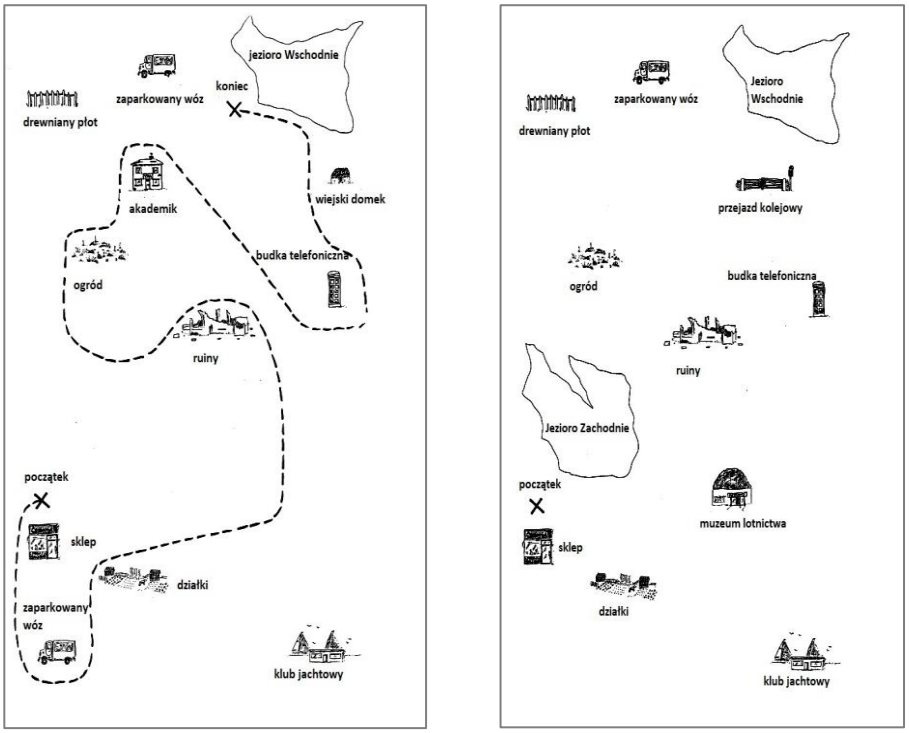

Figure 1
Maps used in the experiment

The follower was supposed to recreate the path as faithfully as possible on the basis of instructions received from the leader. The subjects could use any graphics editor software to draw the path on the map that they received. In this experiment, the task was adapted in such a way that the confederate, a participant who is aware of the experimental task (Branigan and Gibb 2018), always assumed the role of the leader, though the actual participants were informed that the roles were assigned randomly. The participants were randomly assigned to either of two conditions: the i. 'full paragraph' and ii. the 'split utterance' condition, exemplified in Figure 2, meaning that the instructions provided by the confederate were sent in one long message or over several shorter ones.

<MapGuide>	Droga biegnie po lewej stronie sklepu na dół i pod zaparkowanym wozem w prawo. Wóz stoi jakby pod sklepem. 'The road goes down to the left of the shop and to the right under the parked car. The car is parked below the shop.'

<MapGuide>	Droga biegnie na lewo przy sklepie 'The road goes to the left of the shop'
<MapGuide>	i potem w dół 'and then down'
<MapGuide>	i pod zaparkowanym wozem pod sklepem 'and under the parked car under the shop'

Figure 2
A sample from two experimental conditions

The aim of the experiment was to determine whether there is a difference with respect to task success between the two conditions. Therefore, task success was operationalized as a faithful representation of the leader's map by the follower. To compare the degree to which the paths on the maps were faithfully reproduced, each map was cleared from landmarks, so that only the line drawn by the participant remained on the image. The leader's map underwent the same procedure. The pictures were then transferred to ImageMagick software (The ImageMagick Development Team 2021), where they were compared against the leader's map, and the value of mean absolute error in pixels was obtained. In addition, the influence of two other variables – time and tokens – were taken into consideration. The former is the total duration of a conversation, whereas the latter is the raw count of words produced by the leader. This value was obtained by tokenizing the leader's utterances with the use of Natural Language Toolkit (Bird et al. 2009), which is itself a Python programming language library (van Rossum and Drake 2009).

7.2. Data

In total, 22 subjects, who were adults and native speakers of Polish, participated in the experiment. The participants were randomly assigned to either experimental condition (split or full utterances). Out of the twenty-two conversations, two conversations had to be excluded from analysis. The first of the two trials was removed because the data file was corrupted, and the second because of logistic problems during the experiment. Each participant was allocated 30 minutes for the completion of the task, but the actual duration of the conversation varied. In fact, conversations lasted between ~ 14 and ~ 34 minutes (mean = ~ 25).

7.3. Analysis

All statistical analyses were performed in the R programming language (R Core Team 2021). The scores were then tested for normality of distribution with the Shapiro-Wilk test to determine the most optimal statistical test. The normality test proved that the distribution is Gaussian ($p > 0.05$). Therefore, the ANOVA test could be and was performed for the data. In addition to treatment, the predictor variables also included time (the duration of a conversation) and tokens (the number of running words produced by the leader). The results are shown in Table 3.

Table 3
The results of ANOVA test from the experiment

	Df	Sum Sq	Mean Sq	F value	Pr(>f)
Treatment	1	0.0000857	8.573e-05	1.904	0.187
Tokens	1	0.0000037	3.710e-06	0.082	0.778
Time	1	0.0000373	3.730e-05	0.828	0.376
Residuals	16	0.0007205	4.503e-05		

All predictor variables produce an insignificant effect on the outcome variable ($p > 0.05$). This suggests that the mean values of the outcome variable do not differ significantly between the two groups.

7.4. Discussion

The results of the statistical test suggest that the predictor variables do not affect the mean absolute error The results imply that utterance splitting does not determine task success: specifically, participants in each condition reproduced the maps equally faithfully because the values of the mean absolute error in pixels did not differ significantly between the two treatments. The same applies to the duration of conversation and the number of tokens produced by the leader: each predictor variable did not influence task success significantly. On the one hand, it might seem that utterance splitting does not perform the role

that this paper ascribes to it; on the other, humans can communicate successfully even without sharing a linguistic code: if a tourist visiting a foreign country does not speak the local language, he or she can ask for directions using pantomime. Such a communication can be quite successful, but definitely not effortless, as the costs of formulation, production, perception and comprehension spike. Hence, a post-hoc analysis was conducted in order to determine differences in effort between the two conditions.

8. Post-hoc analysis

The aim of the post-hoc analysis was to determine the effort required to achieve task success in both conditions.

8.1. Data and analysis

The data from the experiment was re-analysed. To achieve the aims of the post-hoc analysis, we performed two new ANOVA tests. The test involved one independent variable – the treatment – and two dependent variables: time (the total duration of the conversations) and tokens (the number of running words produced by the leader). In the first step, the difference in time between treatments was assessed (see Table 4). Prior to this, the data was tested for normal distribution with the Shapiro-Wilk test, which showed that the distribution is normal ($p > 0.05$).

Table 4

The results of ANOVA test, time by treatment

	Df	Sum Sq	Mean Sq	F value	Pr(>f)
Treatment	1	239.2	239.15	7.321	0.0145
Residuals	18	588.0	32.67		

In contrast with the previous test, the result is significant ($p < 0.05$). A post-hoc Tukey's HSD test was conducted to determine how duration interacted with condition. The difference of means (by subtracting the duration of the paragraph treatment

from the split utterances treatment) is 6.96, which means that conversations with turns in chunks are, on average, shorter. The results are displayed in Figure 3.

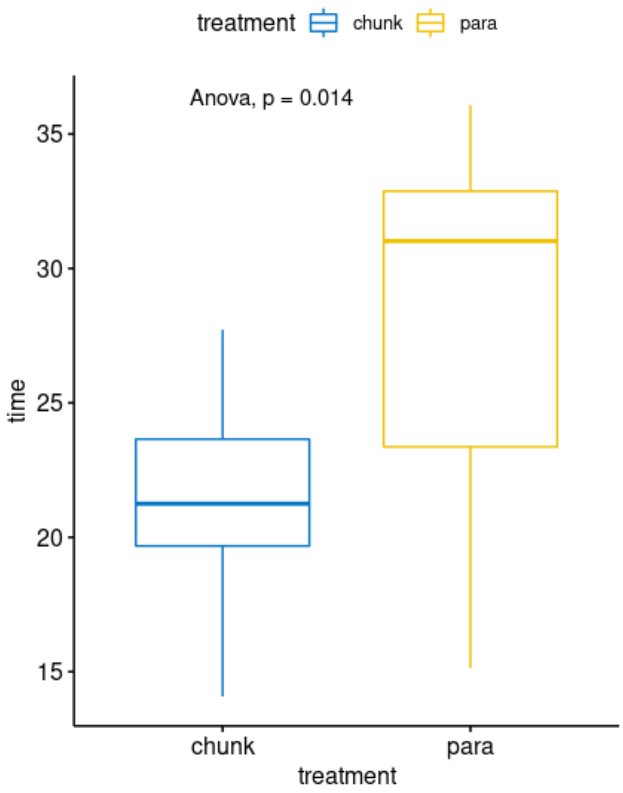

Figure 3
A comparison of total time between conditions

Finally, the number of input instructions produced by the instructor was subject to analysis. The number of instructions was quantified as the total number of tokens produced by the leader. Similarly to the previous two analyses, the data was tested for normal distribution. The Shapiro-Wilk test revealed that the distribution of the number of words per conversation is normal ($p > 0.05$). Therefore, one-way ANOVA test was conducted to verify whether there is a difference between the two

treatments and the number of tokens issued by the instructor. Table 5 contains the results of the analysis.

Table 5

The results of ANOVA test, tokens by treatment

	Df	Sum Sq	Mean Sq	F value	Pr(>f)
Treatment	1	134797	134797	10.5	0.00454
Residuals	18	231096	12839		

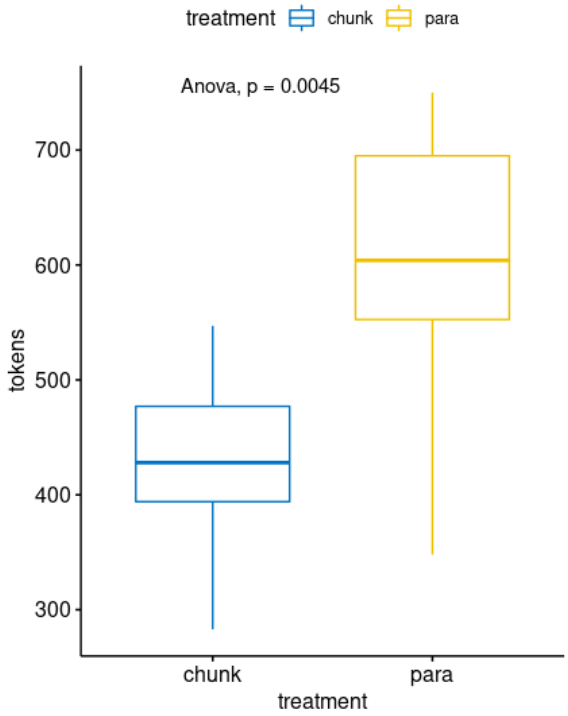

Figure 4

A comparison of the number of tokens produced by the leader between conditions

Similarly to Time by Treatment, this relation is also significant ($p < 0.05$). Again, a post-hoc Tukey's HSD test was conducted to determine how the amount of input produced by the leader (measured in tokens) interacted with the condition. The difference of means (by subtracting the tokens of the paragraph treatment from the tokens from the split utterance) is 165.0202, which means that the leader, on average, needed to produce less input to achieve the same end.

8.2. Discussion

The first analysis determined that successful communication is not driven by the turn-taking strategy of the leader. Notwithstanding the condition, experimental subjects performed equally well – the mean absolute error between the two groups was insignificant. The original analysis also revealed that the time spent on the activity or the number of words produced by the leader do not affect the outcome variable, either. However, which of the strategies is more efficient in QS-CMC was addressed afterwards, in the post-hoc analysis.

The analysis revealed a clear difference between the two groups in terms of communicative effort. The split utterance condition, overall, was more efficient in terms of input (fewer tokens produced by the leader) and duration (conversations were shorter). These results suggest that splitting an utterance may be a strategy used by interactants in QS-CMC in order to trade off the limitations of the medium, communicate efficiently, and thus adhere to the principle of least communicative effort. But what are the reasons that make split utterances more efficient in comparison with full paragraphs?

The first explanation might be that it facilitates the coordination of the content of interaction. Thereby, interactants in QS-CMC imitate face-to-face communication. In face-to-face conversation, the addressee receives the input as it is being produced, 'online', by the current speaker. As a result, the continuous flow of information ensures that new information can be

integrated in real time. This aspect of turn-taking has recently been considered by psycholinguistic research (see Levinson and Torreira 2015). Some researchers in this field, notably Pickering and Garrod (2013a, 2013b), emphasize the role of prediction in language production and comprehension. In dialogue, owing to the so-called forward models, people are supposed to construct perception and action representations during production and comprehension in order to predict the incoming verbal stimulus or to predict their own future actions. The construction of forward models allows for a more rapid exchange of information and the fluency of dialogue. In the case of the experiment, sending messages piece by piece, instead of as a whole paragraph, means that the forward model can be built up earlier, and its assumptions reassessed as new input arrives.

Another content-related explanation can be the distribution of information in dialogue. How much information a given utterance holds can be explained under informational entropy, first proposed by Shannon (1948). According to Shannon (1948), communication occurs over a noisy channel, i.e. the intended message may be corrupted by production or perception errors. The entropy of a random variable is the average level of information that the variable carries. We speak of the highest value of entropy when all interpretations of a variable are equally possible (0.5), and the lowest when one interpretation is completely unlikely (0) and the other is the only possible interpretation (1). Under this approach, information transmitted via a channel with limited bandwidth should be most efficient if entropy levels are distributed uniformly and close to the channel's full capacity (Genzel & Charniak 2002). In fact, research shows that information in written texts and in dialogue is distributed uniformly in the signal (Jaeger 2010, Jaeger & Tily 2011, Qian & Jaeger 2012, Xu & Reitter 2016). In Jaeger's (2010) view, efficient communication balances between communicating the optimal amount of information and sending too much information. Linguistic communication is therefore optimal, if, on average, each word carries the same amount of information, and

the rate at which information is sent is close to the channel's maximum bandwidth. In the case of the experiment reported in sections 6 and 7, although the analysis cannot directly account for information content of words used by the leader, we might argue that utterance splitting is a strategy that transmits linguistic information at an *optimal* rate. Longer messages transmitted at once carry too much information, which influences the overall length of conversation.

Finally, utterance splitting may facilitate coordinating on the procedure in conversation. As Clark and Brennan (1991) argue, interactants not only coordinate the content (what they talk about) of communication, but also on the procedure (synchronisation of actions) in reaching the common ground. Shorter turns may facilitate coordination in such a way that (1) the addressee can better predict turn endings by receiving information at a more-or-less constant rate, (2) if a repair is necessary, it can be initiated almost instantly, and (3) a constant stream of messages ensures keeping the floor by the current writer and, as a result, lowers the chance that the intended addressee introduces a new thread, or a question related to prior instruction. In consequence, utterance splitting means that communication consumes less effort because less time has to be spent on maintaining and controlling the turn-taking procedure.

9. Conclusions

The aim of the paper was to overview theories of turn-taking in quasi-synchronous text-based communication and to experimentally test the influence of turn-taking strategies, especially utterance splitting, in text-based chat. Three theories of turn-taking were presented, but none of them provided a sufficient explanation of choosing a strategy. The experiment reported here approached utterance splitting from the perspective of the principle of least cooperative effort, which posits that interactants try to lower the effort they invest in communication. The results indicate that although both split utterance and full

paragraph lead to communicative success, the former is much more efficient. These outcomes were discussed in the light of facilitating the construction of forward models, uniform distribution of information in the channel, and coordinating on turn-taking as possible explanations of greater efficiency of utterance splitting.

Information about funding

This research was supported by the National Science Centre (NCN) under grant agreement UMO-2019/33/N/HS2/00541.
Praca powstała w wyniku realizacji projektu badawczego o nr UMO-2019/33/N/HS2/00541 finansowanego ze środków Narodowego Centrum Nauki.

References

Anderson, Anne H., Miles Bader, Ellen G. Bard, Elizabeth Boyle, Gwyneth Doherty, Simon Garrod, Stephen Isard, Jacqueline Kowtko, Jan McAllister, Jim Miller, Catherine Sotillo, Henry S. Thompson, Regina Weinert (1991). "The Hcrc map task corpus". *Language and Speech* 34/4: 351–366.

Anderson, Jeffrey, Fred Beard, Joseph Walther (2010). "Turn-taking and the local management of conversation in a highly simultaneous computer-mediated communication system". *Language@Internet* 7. Available at <https://www.languageatinternet.org/articles/2010/2804>. Accessed 20.10.2021.

Baron, Naomi S. (2010). "Discourse structures in instant messaging: the case of utterance breaks". *Language@Internet* 7/4. Available at <http://nbn-resolving.de/urn:nbn:de:0009-7-26514>. Accessed 20.10.2021.

Bavelas, Janet, Linda Coates, Trudy Johnson (2002). "Listener responses as a collaborative process: the role of gaze". *Journal of Communication* 52: 566–580.

Beattie, Geoffrey W., Anne Cutler, Mark Pearson (1982). "Why is Mrs Thatcher interrupted so often?" *Nature* 300: 744–747.

Bird, Steve, Ewan Klein, Edward Loper (2009). *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*. Sebastpol: O'Reilly.

Branigan, H. P., C. L. Gibb (2018). "Structural priming". In: Annette M. B. de Groot, Peter Hagoort (eds.). *Research Methods in Psycholinguistics and the Neurobiology of Language: A Practical Guide*. Oxford: Wiley, 130–150.

Chafe, Wallace L. (1987). "Cognitive constraints on information flow". In: Russell S. Tomlin (ed.). *Coherence and Grounding in Discourse*. Amsterdam: John Benjamins Publishing Company, 21–51.

Chafe, Wallace L. (1980). "The deployment of consciousness in the construction of narrative". In: Wallace L. Chafe (ed.). *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. Norwood, NJ: Ablex, 9–50.

Clark, Herbert H. (1996). *Using Language*. Cambridge: Cambridge University Press.

Clark, Herbert H., Susan E. Brennan (1991). "Grounding in communication". In: L. B. Resnick, J. M. Levine, S. D. Teasley (eds.). *Perspectives on Socially Shared Cognition*. Washington: American Psychological Association, 127–149.

Clark, Herbert H., Deanna Wilkes-Gibbs (1986). "Referring as a collaborative process". *Cognition* 22/1: 1–39.

Cutler, Anne, Mark Pearson (1986). "On the analysis of prosodic turn-taking cues". In: Catherine Johns-Lewis (ed.). *Intonation in Discourse*. London: Routledge.

Ford, Cecilia, Sandra Thompson (1996). "Interactional units in conversation: syntactic, intonational, and pragmatic resources for the management of turns". In: Elinor Ochs, Emanuel A. Schegloff, Sandra A. Thompson (eds.). *Interaction and Grammar*. Cambridge: Cambridge University Press, 134–184.

Gambi, Chiara, Torsten K. Jachmann, Maria Staudte (2015). "Listen, look, go! The role of prosody and gaze in turn-end anticipation". In: *Proceedings of the 37th Annual Meeting of the Cognitive Science Society, CogSci 2015, Pasadena, California, USA, July 22-25, 2015.*

Garcia, Angela, Jennifer B. Jacobs (1998). "The interactional organization of computer mediated communication in the college classroom". *Qualitative Sociology* 21/3: 299–317.

Garcia, Angela, Jennifer B. Jacobs (1999). "The eyes of the beholder: understanding the turn-taking system in quasi-synchronous

computer-mediated communication". *Research on Language and Social Interaction* 32/4: 337–367.

Genzel, Dmitriy, Eugene Charniak (2002). "Entropy rate constancy in text". *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL'02)*, 199–206.

Gibson, Will J. (2014). "Sequential order in multimodal discourse: Talk and text in online educational interaction". *Discourse & Communication* 8/1: 63–83.

Grice, H. Paul (1975). "Logic and conversation". In: Peter Cole, Jerry Morgan (eds.). *Speech Acts* [*Syntax and Semantics* 3]. New York: Academic Press, 41–58.

Herring, Susan (1999). "Interactional coherence in CMC". *Journal of Computer-Mediated Communication* 4/4: Article 4.

Jaeger, T. Florian (2010). "Redundancy and reduction: Speakers manage syntactic information density". *Cognitive Psychology* 61/1: 23–62.

Jaeger, T. Florian, Harry J. Tily (2011). "On language 'utility': Processing complexity and communicative efficiency". *Wiley Interdisciplinary Reviews. Cognitive Science* 2/3: 323–335.

Kendon, Adam (1967). "Some functions of gaze-direction in social interaction". *Acta Psychologica* 26: 22–63.

Levinson, Stephen C., Francisco Torreira (2015). "Timing in turn-taking and its implications for processing models of language". *Frontiers in Psychology* 6: 731.

Oliveira, Miguel, Tiago Freitas (2008). "Intonation as a cue to turn management in telephone and face-to-face interactions". *Speech Prosody 2008*: 485–488.

Pickering, Martin J., Simon Garrod (2004). "Toward a mechanistic psychology of dialogue". *Behavioral and Brain Sciences* 27/2: 169–190.

Pickering, Martin J., Simon Garrod (2013a). "An integrated theory of language production and comprehension". *Behavioral and Brain Sciences* 36/4: 329–347.

Pickering, Martin J., Simon Garrod (2013b). "Forward models and their implications for production, comprehension, and dialogue". *Behavioral and Brain Sciences* 36/4: 377–392.

Qian, Ting, T. Florian Jaeger (2012). "Cue effectiveness in communicatively efficient discourse production". *Cognitive Science* 36/7: 1312–1336.

Riordan, Monica A., Kris M. Markman, Craig O. Stewart (2013). "Communication accommodation in instant messaging: An examination of temporal convergence". *Journal of Language and Social Psychology* 32/1: 84–95.

Rossano, Frederico (2012). "Gaze in conversation". In: Jack Sidnell, Tanya Stivers (eds.). *The Handbook of Conversation Analysis*. John Wiley Sons, Ltd., 308-329.

Rühlemann, Christoph, Stefan T. Gries (2020). "Speakers advance-project turn completion by slowing down: A multifactorial corpus analysis". *Journal of Phonetics* 80: 100976.

Sacks, Harvey, Emanuel A. Schegloff, Gail Jefferson (1974). "A simple systematic for the organisation of turn taking in conversation". *Language* 50: 696–735.

Schaffer, Deborah (1983). "The role of intonation as a cue to turn taking in conversation". *Journal of Phonetics* 11/3: 243–257.

Schönfeldt, Juliane, Andrea Golato (2003). "Repair in chats: a conversation analytic approach". *Research on Language and Social Interaction* 36: 241–284.

Shannon, C. E. (1948). "A mathematical theory of communication". *Bell System Technical Journal* 27/3: 379–423.

Simpson, James (2005). "Conversational floors in synchronous text-based CMC discourse". *Discourse Studies* 7/3: 337–361.

The ImageMagick Development Team (2021). *ImageMagick* (7.0.10) [Computer software]. Available at <https://imagemagick.org>. Accessed 01.10.2021.

Thompson, Henry S., Anne Anderson, Ellen G. Bard, Gwyneth Doherty-Sneddon, Alison Newlands, Cathy Sotillo (1993). "The HCRC map task corpus: natural dialogue for speech recognition". *Human Language Technology: Proceedings of a Workshop*. New Jersey.

Tudini, Vincenza (2015). "Extending prior posts in dyadic online text chat". *Discourse Processes* 52/8: 642–669.

Uthus, David C., David W. Aha (2013). "The Ubuntu chat corpus for multiparticipant chat analysis". In: *AAAI Spring Symposium – Technical Report*, 99–102.

Van Rossum, Guido, Fred L. Drake (2009). *Python 3 Reference Manual*. Scotts Valley: CreateSpace.

Xu, Yang, David Reitter (2016). "Entropy converges between dialogue participants: explanations from an information-theoretic perspective". In: *Proceedings of the 54th Annual Meeting of the*

Association for Computational Linguistics, Volume 1: Long Papers, 537–546.

Zemel, Alan (2005). Texts-in-interaction: collaborative problem-solving in quasi-synchronous computer-mediated communication". In: *Computer Supported Collaborative Learning 2005: The Next 10 Years!*, 753–757.

Zlatev, Jordan, Sławomir Wacewicz, Przemysław Żywiczyński, Joost van de Weijer (2017). "Multimodal-first or pantomime-first? Communicating events through pantomime with and without vocalization". *Interaction Studies* 18: 454–477.

Żywiczyński, Przemysław, Sławomir Wacewicz, Sylwester Orzechowski (2017). "Adaptors and the turn-taking mechanism: the distribution of adaptors relative to turn borders in dyadic conversation". *Interaction Studies* 18: 276–298.

Marek Placiński
ORCID iD: 0000-0001-7225-3099
Wydział Humanistyczny
Instytut Językoznawstwa
Uniwersytet Mikołaja Kopernika w Toruniu
Fosa Staromiejska 3
87-100 Toruń
Poland
marpla@umk.pl