Speak *now* or forever hold your peace: power law chronemics of turn-taking and response in asynchronous CMC

Yoram M Kalman

Gilad Ravid

Daphne R Raban

Sheizaf Rafaeli

Abstract

Turn-taking is a key characteristic of conversation, and is well studied in diverse contexts. How do insights about chronemics in face-to-face communication transfer to the online context? The research reported here looks into turn-taking in asynchronous computer mediated communication by analyzing three diverse datasets comprising a total of more than 150,000 responses: email responses created by corporate employees in work contexts, responses created by university students participating in course discussion groups, and responses to questions posted in a public, commercial information seeking virtual forum. The response times revealed in each of the datasets yielded a power-law distribution. Moreover, most of the responses (at least 70%) were created within the average response time of the responders, and very few (at most 4%) of the responses were created after a period longer than ten times the average response time. These generalized findings persist across various users, contexts, and even across traditional as well as online communication media. Asynchronous CMC seems more synchronous than is implied by its reputation. email implies a widespread expectation of fairly rapid response. The implications of this uniformity are discussed, as well as some possible applications of the findings to the improvement of computer mediated communication and to the research of online silence.

Introduction

Conversational rhythms, turns and gaps

Conversations are rhythmic in nature, and the rhythms of conversation have long attracted the attention of diverse researchers (e.g. Brady, 1965; Cappella, 1979; Jaffe & Feldstein, 1970; Sacks, Schegloff, & Jefferson, 1978). The on-off pattern determines the rhythm of the conversation, and the pauses or gaps in speech that constitute that pattern, were investigated in depth under various names, including pause, gap and silence (McLaughlin, 1984). The temporal patterns of spoken conversation were researched by digitizing the vocal patterns of monologues and dialogues, and measuring the lengths of conversational categories such as a vocalization, a pause, a switching pause, floor times, etc. An interesting generalization emerging from this literature is that when the lengths of each of these classification categories are plotted on a semi-logarithmic graph the points tend to fall along a straight line, a phenomenon that has "… generally been found to be exponential." (Jaffe & Feldstein, 1970), page 25. This exponential relationship is a manifestation of skewed distributions, in which the majority of the pauses are very short, and only a minority is of average or above length. It is interesting to note that when the results of Jaffe and Feldstein are extracted from the original graphs and re-plotted using contemporary computerized statistical tools, the power law distribution seems to yield an even better fit than the exponential distribution. For details see the discussion section below.

Turn-taking

Sacks et al.(1978) suggested a model to explain the relatively rapid turn-taking transitions, as well as many other aspects of turn-taking in naturally occurring spoken conversation, emphasizing that "the presence of 'turns' suggests an economy, with turns for something being valued…" (page 7). Their model is based on a set of rules *"…providing for the allocation of the next turn to one party, and coordinating transfer so as to minimize gap and overlap. For any turn:*

1. *At initial turn-constructional unit's initial transition-relevance place:*

a. *If the turn-so-far is so constructed as to involve the use of a 'current speaker selects next' technique, then the party so selected has rights, and is obliged, to take next turn to speak, and no others have such rights or obligations, transfer occurring at that place.*

b. *If the turn-so-far is so constructed as not to involve the use of a 'current speaker selects next' technique, self-selection for next speakership may, but need not, be instituted, with first starter acquiring rights to a turn, transfer occurring at that place.*

c. *If the turn-so-far is so constructed as not to involve the use of a 'current speaker selects next' technique, then current speaker may, but need not, continue, unless another self selects.*

2. *If, at initial turn-constructional unit's initial transition-relevance place, neither 1(a) nor 1(b) has operated, and, following the provision of 1(c), current speaker has continued, then the Rule-set (a)-(c) reapplies at next transition-relevance place, until transfer is affected."* (page 13)

A key element of naturally occurring conversation is the presence of conversational gaps, explained by Sacks et al. to reflect the optional nature of turn-taking under rules 1(b) and 1(c). The gaps are described (McLaughlin, 1984) as (1) hesitation pauses, (2) switching pauses or (3) initiative time latencies. A hesitation pause is a within-turn pause by the speaker, while a switching pause is a within-turn pause at which the floor is switched between one speaker and another. Initiative time latencies describe a gap in the speech of an individual speaker who realizes that despite her or his expectation for a response, the other side remains silent. McLaughlin & Cody (1982) define a conversational lapse as "an extended silence (3 seconds or more) at a transition-relevance place, in a dyadic encounter the focus of which is conversation". They go on to define some obvious exceptions to this definition, and to explain the 3-second of silence as an "awkwardness limen". Recent neurological work corroborates this finding, defining the range of 2-3 seconds as a temporal integration time range which is a general principle of the neurocognitive machinery (Poppel, 2004; Vollarth, Kazenwadel, & Kruger, 1992).

Interactivity

The organization of language is a result of an *interactive* process between the participants in linguistic interaction, and "Rather than simply producing language and other semiotic structure, participants in interaction are attributing complex cognitive and inferential practices to their coparticipants and taking these into account in the detailed organization of ongoing social action" (Goodwin, 2002). Interactivity refers to the extent to which communication reflects back on itself, feeds on and responds to the past. Interactivity is the degree of mutuality and reciprocation present in a communication setting. The term interactivity is widely used to refer to the way content expresses contact, and communication evolves into community. And, interactivity is a major option in governing the relation between humans and computers (Rafaeli, 1984; Rafaeli, 1988; Rafaeli, 2004). Interactivity is an essential characteristic of effective online communication, and has an important role in keeping message threads and their authors together. Interactive communication (online as well as in more traditional settings) is engaging, and loss of interactivity will result in a breakdown of the communicative process. Research of rhythms in email and other Computer Mediated Communication (CMC) media resulted in claims that text-only CMC is "Interactionally Incoherent": disjointed, without clear turns, and in general "chaotic". But, as noted by Herring (1999), text-only CMC is extremely popular, despite obstacles such as disrupted turn adjacency and lack of simultaneous feedback. The online interaction is highly desired, and almost addictive in nature (Caplan, 2003; Morahan-Martin & Schumacher, 2000) despite the apparent incoherence. That leads Herring to claim that unique attributes of CMC are actually leveraged by users to intensify interactivity and extend the limits of traditional, spoken, conversation.

Online responsiveness

Turn taking and interactivity are closely linked. Failure to respond or to take the floor creates a breakdown of interactivity. Online interactivity and responsiveness were carefully studied in various contexts: responsiveness and response time to customers who email an organization or post an online inquiry (e.g. Customer-Respect-Group, 2004; Hirsh, 2002; Mattila & Mount, 2003; Stellin, 2003; Strauss & Hill, 2001); responses to online surveys (e.g. Lewis, Thompson, Wuensch, Grossnickle, & Cope, in press; Sheehan & McMillan, 1999); responsiveness to business correspondence (e.g. Abbott et al., 2002; Pitkin & Burmeister, 2002; Tyler & Tang, 2003); and, work on response times in discussions on Usenet (Jones, Ravid, & Rafaeli, 2004) and to questions posted to the “Google Answers” website (Edelman, 2004; Raban, Ravid, & Rafaeli, 2005). An examination of these reports reveals a recurring pattern that closely resembles the findings on conversational pauses, described above: most of the responses were created within relatively short times, and only a minority of them is of average or above length. This chronemic (Walther & Tidwell, 1995) distribution was described in detail in work carried out on the responsiveness profile of email-using Enron employees (Kalman & Rafaeli, 2005). In that work, the “windfall” of the confiscation and release of massive data files of the Enron Corporation, made it possible to extract detailed behavioral information without raising privacy and other ethical limitations. The observed pattern in the Enron corpus was a concentration of most of the responses within a relatively short period of time, and a spread of ever increasing response times at a relatively very low frequency. When plotting the frequency of responses against response latency, the resultant distribution is highly skewed to the left, with a stretched out and rapidly diminishing thin right tail. The results, which are more robust quantitatively than any previous work on online responsiveness, clearly reveal the distribution patterns of pauses already observed in traditional conversation, as well as in online communication. This apparent similarity between traditional spoken conversation, and online, persistent conversation (Erickson, 1999; Erickson & Herring, 2005) triggered this work in which we set out to generalize these initial findings from the Enron Corpus, by analyzing additional datasets, and looking for properties common to these online conversations, as well as for properties shared by online conversation and traditional spoken conversation.

Power Law distributions

The power law distribution has been observed in many fields and in diverse phenomena (Axtell, 2001; Comellas & Gago, 2005; Gabaix, Gopikrishanan, Pelrou, & Stanley, 2004; Keeling & Grenfell, 1999; Qian, Luscombe, & Gerstein, 2001; Reed, 2001; Zipf, 1949). Two famous power law distributions are the Pareto distribution and Zipf’s law. The power law is also similar to the lognormal distribution. The graphical manifestation of a power law distribution is a straight line on a log-log graph, since it is an expression of the relationship $y=ax^b$. The similarities between phenomena that are so diverse in nature is a source for confusion as well as for innovative modeling in efforts to identify common underlying mechanisms that lead to power law distributions or to distributions which are similar to it (Adamic, 2005; Goldberg, Franklin, & Roth, 2005; Mitzenmacher, 2003).

The research question

This research explores the question whether persistent conversation shares fundamental properties with traditional, spoken conversation. Specifically, our research question was:

Are chronemic distribution patterns similar for turn-taking pauses in spoken and persistent conversation? If so, what are these common patterns?

The datasets of turn-taking pauses analyzed here originated from conversations carried out via diverse technologies, within a variety of contexts and across different individuals and groups of individuals.

Methodology

Characterizing aggregate and individual response times

Three distinct datasets of asynchronous computer mediated communication were analyzed. The first dataset, "Enron emails" includes the response times of corporate email users. Data are extracted from the correspondence of Enron employees and is described in detail in Kalman and Rafaeli (2005). Since the response times in this dataset are based on subtractions between time stamps created by two separate computers, some negative response times were also recorded. In the present analysis 15,815 registered response times (positive and negative) were used with the exception of seven outliers assumed to be measurement errors: two extremely high, and five extremely low and negative results. The second dataset, "University forum" is described in Ravid & Rafaeli (2004). It contains 115,416 response times to messages posted by thousands of university students to online, course related, asynchronous discussion groups during 1999-2002. The third dataset, "Google answers" is described in Raban, Ravid et al. (2005). It contains 40,072 response times of answers to questions posted to Google answers (http://answers.google.com). Each of these three aggregate response time datasets was analyzed separately by the same methods used for identifying power law distributions, as well as response latencies in traditional spoken communication: the response times were grouped into bins and plotted on a log-log graph; regression analysis was performed for the power distribution, and a coefficient of determination (R^2) was calculated; Various binning methods and truncation possibilities were tried, to refine the presentation, alterations which did not materially affect the R^2 of the regression analysis. The percentiles of the average response time as well as the percentiles of ten times (10x) that average response time were calculated. The percentile analysis was then repeated for all individual users in the "Enron emails" dataset, as well as for 15 individuals from the "Google answers" dataset: the five users with the largest number of responses, as well as ten of the users with 100-120 responses. Finally, a small sample of "Enron emails" responses that were created after a long delay were selected and their content inspected.

Results

Distribution of response times

115,416 University forum responses, 15,815 Enron email responses and 40,072 Google answers were analyzed. Despite the diverse sources of these responsiveness profiles, when plotted on a log-log graph, all three datasets presented a power-law distribution (Figure 1 a-c).

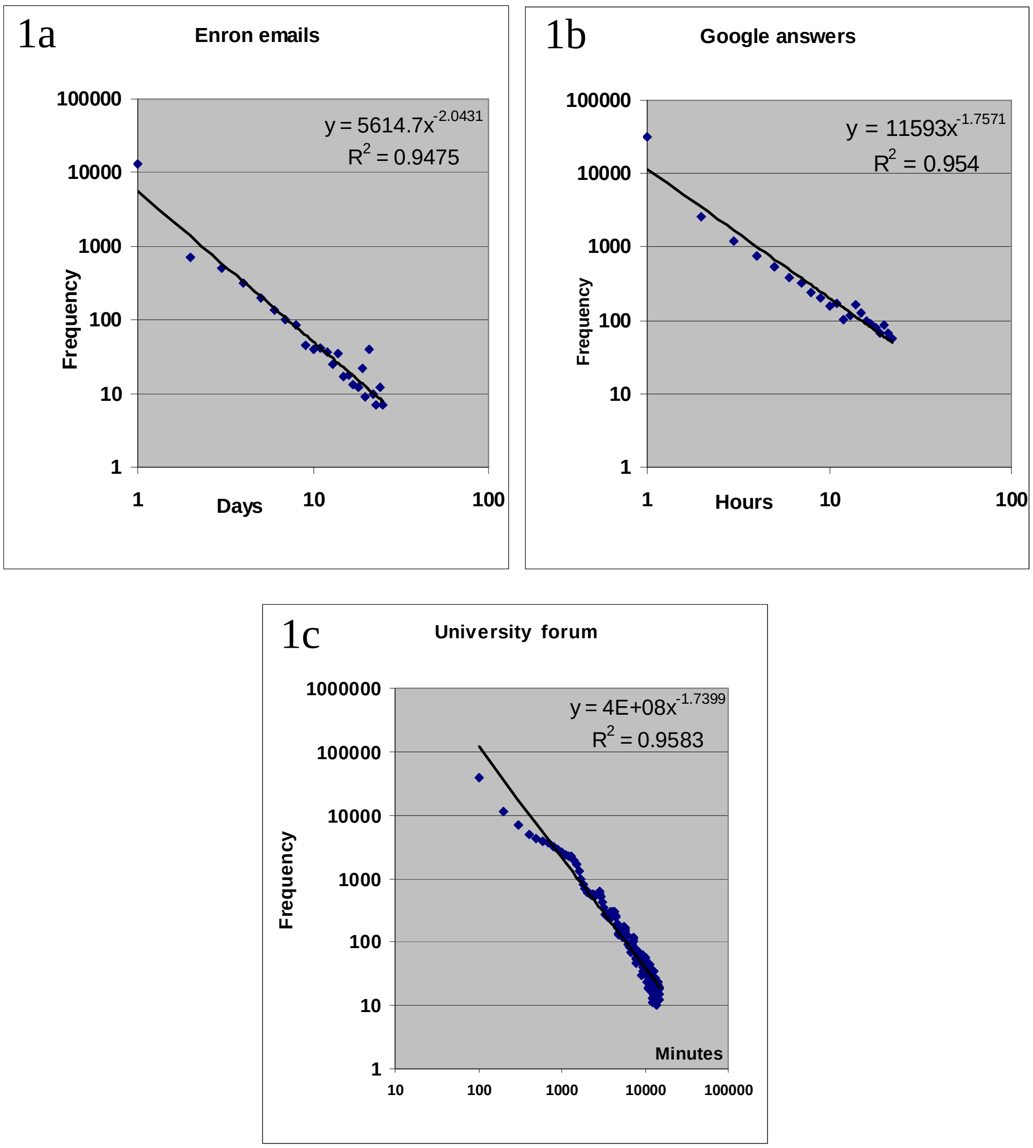

Figure 1: *power law plots of the cumulative response times of the three datasets.*

An analysis of the distribution of each of the datasets revealed that the average response time in each of the datasets falls at or above the 80th percentile. It also revealed that ten times that average time in each of the datasets falls at or above the 97th percentile (Table 1).

Dataset	**Average response time**	**Percentile rank of average response time**	**Percentile rank of 10x average response time**
Enron emails	28.76 hours	86%	97%
University forum	23.52 hours	80%	99%
Google answers	1.58 hours	84%	97%

Table 1: *Average response times in each dataset, and the percentile rank of that average response time, and of ten times (10x) that average response time, for each dataset.*

This remarkable similarity between datasets comprising aggregate responses created under such diverse circumstances and by many individuals elicited the question whether this generalization about percentiles is a result of the aggregation of many response times, or if it is also reflected in the behavior of individual users. An analysis of the 74 Enron email users for whom more than 50 unique responses existed showed that only 65% of them (48) met the strict criteria that their average response time was at or above the 80th percentile. But, a slight relaxation of the criteria revealed that 95% of them (70) created 70% or more of their responses within less than their average response time. And, of these 74 users, only five users' 10x of average response time was below the 97th percentile, and none were below the 96th percentile. The 15 users from the Google answers database displayed a similar behavior: 93% (14) created more than 70% of their responses within their average response time or less, and all of them created at least 96%

of their responses with less than 10x their average response time. In summary, the vast majority of the individual users created most (70% or more) of their responses within their average response time; and, almost all (96% or more) of their responses within a span of time equal to ten times their average response time. This relaxed generalization also holds for the cumulative results.

Discussion

The results of the distribution analyses are that all three user groups show, in aggregate, a similar power law distribution. A closer inspection of the distributions shows that despite the significant differences between the types, purpose and context of asynchronous conversations taking place within each group, in all three of them at least 80% of the responses are created within the average response time of that group, and at least 97% of the responses are created within ten times that average response time. In cases where analysis was possible, even individual users show the same skew: at least 70% of almost each individual’s responses were made within that user’s average response time, and at least 96% within ten times his or her average response time.

The generalizability of the findings

The findings point to common chronemic characteristics of CMC. The three datasets described are very diverse in their characteristics: they represent diverse populations (business people, students, and varied internet users in a public arena), assorted asynchronous text based CMC technologies (email, discussion forum, web pages), a variety of contexts (academic education, major corporation, competitive online bidding), a range of average response times (from 1.5 hours to a little over a day) and of cohort sizes (more than 15,000 to more than 100,000, a total of over 170,000 responses), a period spanning at least seven years, and respondents from the US as well as from other countries. Despite these differences, a recurring pattern surfaces when analyzing the aggregates: a power law distribution of the response times that can be described by the generalization that regardless of the average response time, most (at least 80%) of the responses are already created within that average time, and almost all (at least 97%) of the responses are created within 10x of the average response time.

The strength of this generalization is further revealed when drilling down to the level of individual users. We have shown that the generalizations at the aggregate level need to be only slightly relaxed (from 80% to 70% and from 97% to 96%) in order to describe the vast majority of dozens of users from the two datasets in which personal identification was possible, and users for whom a sufficiently large sample of response times was available. This finding is an indication that users of asynchronous CMC tend to create responses within a relatively short time, in the order of magnitude of the average response time and are unlikely to respond after a duration longer than one order of magnitude higher than that average response time.

The quantitative findings described here allow quantifying the probabilities of response events, based on estimated or measured average response times. The practical implications of these findings lie for example in the potential to increase social translucence in online communication. Social translucence is described as a system that makes social information visible and enables participants to be both aware of what is happening, and to be held accountable for their actions as a consequence of public knowledge of that awareness (Erickson & Kellogg, 2000). For example it is relatively simple to construct a tool that will be able to use these quantitative findings to analyze the responsiveness profiles of specific people one is communicating with via email, and estimate the probability of a response from each of them within a specified period of time. The same mechanism can also be applied by users who wish to analyze their own responsiveness profile and use the conclusions of this analysis to improve their responsiveness.

The robustness of the generalization receives further substantiation when one looks at well established rules describing latencies and response times in traditional forms of communication. For example, in Jaffe & Feldstein's work (1970) on face-to-face contexts, the quantitative results of the duration of pauses of one speaker in a face to face dialogue (pp. 76, figure IV-9) present the same characteristics as any of the three datasets described here: 70-80% of the pauses are shorter than the average pause length (estimated at .97 seconds), and a pause of 10x of that average, more than 9.7 seconds, did not occur even once in that 50 minute dialogue. Moreover, when the plot is reconstructed using modern statistical tools, and regression analysis is performed, the power law distribution gives a high R^2 value of .82, even better than that for the exponential distribution reported by the authors (the calculation was not performed by the authors, but the reconstruction of the data by us gives an R^2 of 0.74 for an exponential distribution). Similar behavior appears in phone based conversations such as those described by Brady (1968), though precise analysis is difficult due to partial presentation of the results. Apparently, the power law distribution, as well as the tendency to keep most of the response latencies relatively short, are a universal characteristic of typical human communication, using traditional as well as new technologies. We believe it is now possible to reevaluate and re-analyze not only data produced decades ago, but also more recently published data, to further test the generalizations suggested here. For example, an additional dataset, that originates in an online report (Hamilton, 2005), summarizes response times in 199 online surveys in which 523,790 invitations were sent and over 170,000 responses were received. Though we did not have direct access to the dataset, the report describes a similar pattern to the one observed here, where an estimated 70% of the responses were created within the average response time (a little less than 3 days), and where over 99% of the responses were created by four weeks (10x the average response time). Additional published work in various disciplines apparently reveal behavior which is in agreement with these generalizations (Jones et al., 2004; Matzler, Pechlaner, Abfalter, & Wolf, 2005; Strauss & Hill, 2001). It would be interesting and instructive to find occasions in which the same rules apply, as well as exceptions to the rules. This can be achieved by further analysis of published data, as well as by dedicated original research that will focus on asynchronous CMC, including areas not mentioned here such as, for example, response times within blogs. Furthermore, research should look at response times in synchronous CMC such as instant messaging, chatting and SMS'ing. For a discussion of synchronous vs. asynchronous CMC see Newhagen & Rafaeli (1996).

Possible explanations for the findings

Why do people create most of their responses within a relatively short period? One of the promises of online communication was thought to be its asynchronicity: the ability to respond at one's convenience, even after a relatively long wait (e.g. Lantz, 2003; Newhagen & Rafaeli, 1996). Why then do we see that in practice most responses are created quickly, and that if a response is not created within a short period of time, the probabilities for a response drop precipitously?

One possible answer is the well documented phenomenon of online information overload (Davenport & Beck, 2001; Shenk, 1999): as messages flow in, people either respond to them at once, or put them aside and rarely return to them. Evidence for this behavior was already presented (Jones et al., 2004). Another possible explanation for that behavior pattern is linked to the signaling power of a quick response: in asynchronous CMC, a quick response is one of the only non-verbal tools that can be used to signal immediacy, care and presence. Thus, there is a preference for quick replies (Aragon, 2003; Danchak, Walther, & Swan, 2001; Feldman & March, 1981; Goodwin, 2002; Walther & Tidwell, 1995).

The full explanation for the rapid answers probably lies in the combination of both principles mentioned in the previous paragraph: due to the practical nature of online communication in an age of information overload, a quick response is the best way to ensure that a response will be created. And, by sending a quick response, one conveys rapport, immediacy and presence. Moreover, the practicality of interactive communication is that of immediate responses. It is difficult to imagine a world in which every message, even one that was delivered a long time ago, has a high probability of receiving a response. In this context, it is interesting to compare the actual wording used in texts created within a typical response time, to those created after an unusual delay. This analysis was possible in the Enron emails dataset. A qualitative examination of some of those slow responses, (those created after a relatively long period of time and well beyond average response time) implies that these responses are different from the responses created within the average response time. These responses are more likely to mention the long response time, often apologize about the delay and/or provide an explanation. In addition, these are sometimes not really responses, although they were created by replying to a previous email message. They might show the user sending a reply asking about the progress of an issue mentioned in the original email, or even not connected at all to the text of the original email, possibly as a shortcut to typing an email address. See table 2

Response time	Quoted Text	Category
16 days	sorry for the delay	Apology
14 days	Sorry it has taken me so long to write	Apology
18 days	i got back from almost three weeks vacation yesterday and am back at work	Explanation
14 days	i just got back into town from almost 3 weeks vacation. sorry i didn't get in touch over the holidays, but…	Explanation + apology
23 days	Only took me 3 weeks to respond. That's pretty good for me. I think things started collapsing the day I got your original email	Humorous apology + explanation
16 days	Just following up to see if the recruiting season has started and to make sure everything is going okay. If you need anything, just say	Reference to subject in original email. Not an answer to question
51 days	Dale, how are we coming on this project in relation to the info Esther sent you? Do you need anything else from Esther? Thanks.	Reference to subject in original email. Not an answer to question
109 days	Hey Mom… thought i would give a call but don't have your number at work. send it if you get a chance. love,	email response as probable short-cut to typing email address

Table 2: *examples of texts from messages in the Enron email database (Kalman & Rafaeli, 2005). Examples were taken from responses created after a long delay, beyond 10x the average response time.*

This initial observation should be studied further using formal content analysis, comparing responses created within different percentiles of an individual's responsiveness profile, as well as percentiles of the average responsiveness profile in the user's reference group.

Turn taking rules in asynchronous CMC

If the generalizations about the timing of switching pauses in turn taking online are comparable to those of traditional turn-taking, how do the rules of traditional turn-taking apply to asynchronous CMC? The set of rules suggested by Sacks et al.(1978) were structured to accommodate fourteen facts about any traditional “mouth to ear” conversation (pages 10-40). Most of these conditions do apply to asynchronous CMC, with three important exceptions, that result from the asynchronous nature of the conversation: in asynchronous CMC conditions 2 and 3 (“overwhelmingly, one party talks at a time” and “occurrences of more than one speaker at a time are common, but brief”) do not apply, due to the permanence (Erickson & Herring, 2005) of the conversation. Permanence (or durability) refers to the fact that in CMC the message is available as well as retained longer for further and repeated examination. Durability of messages overcomes the aural and cognitive difficulty of synchronously processing more than one stream of talk, and allows a separation in time between the receipt of the words, and their processing. In other words, while traditional aural communication is of a “half-duplex” nature, and requires back channels in order to convey additional information while speech is created by one side, CMC allows “full-duplex” communication. In addition, rule 4 (“transitions from one turn to a next with no gap and no overlap between them are common. Together with transitions characterized by slight gap or slight overlap, they make up the vast majority of transitions”) needs to be restated in light of the findings reported in this paper. Our proposal for the restatement of conditions 2 and 3 is that “One party or more can create a message at any given time.” For condition 4 the proposed restatement is “the vast majority of the transitions occur within a relatively short time.” The use of the word “relatively” is intentional. It alludes to the relativity reported in this paper, that no matter if the average response time in that specific conversation is a few hours or even days, the majority of the responses are created within that average length of time, and the vast majority of the responses shortly thereafter. It is important to note here that we measured the response time when there *was* a response. If there was no response, no conversation has taken place.

After having restated three of the 14 conditions, how should the rules for turn taking in persistent conversation be restated to fit the revised conditions?

1. At the moment a message is sent by one party (the sender) to one or more parties (the recipients):

   a. If the sender has selected the next speaker, the party so selected has rights, and is obliged, to send a response as soon as is practicable. Other recepients too have the right to send a response
   b. If the sender has not selected the next speaker, each recipient has the right but not the obligation to send a response
   c. The sender may continue with an additional message
2. If, after a message is sent by the sender, either 1(a) 1(b) or 1(c) has operated, each party who created a reply is assigned the role of sender and the rule-set (a)-(c) reapplies.

Unresponsiveness and silence in asynchronous CMC

These findings on responsiveness, interactivity and maintaining conversation threads in CMC, provide us with tools to investigate the instances when unresponsiveness and silence disrupt the conversation. Despite the extensive research on matters relating to silence in traditional settings, little research on this topic was carried out in online settings: Anecdotal evidence of the need to acknowledge silence as a factor in human-computer communication is nicely described by Nicholas Negroponte from MIT's Media Lab in work that was carried out as early as1978 (Negroponte, 1994); Lurking, a special form of online silence, is being researched (Nonnecke & Preece, 2000; Rafaeli, Ravid, & Soroka, 2004); and, Cramton (2001) documented the disruptive effect silence can have on teams attempting to collaborate online. One of the factors limiting research of online silence is the lack of a basis for the definition of the length of unresponsiveness that constitutes online silence, such as the three second or more "conversational lapse" described above (McLaughlin & Cody, 1982) . The results reported here allow a quantitative definition of online silence as "no response after a period of more than ten times of the average response time". This definition gives a confidence level of at least 95% that a response will no longer be created since our findings show that only 3-4% of the responses are created after that time. It is important to note that this definition is context sensitive, since the average response time is context sensitive. We believe the "above 10x average response time" definition to be conservative, mainly since response rates are usually less than 100%. Moreover, at least a minority of the very late responses created seems not to include actual answers to the original message.

The strength of this definition of a “CMC lapse” is that it combines the rigor of a quantitative, statistical definition, with the ability to adjust for qualitative differences between datasets through its context sensitivity. Thus, when researching online silence in a specific context, researchers will identify an average response time relevant for the context of that specific research. Once that average response time is identified (through the analysis of a large enough dataset, or through the use of a relevant published average), it can be assumed that if a response was not created within that 10x period of time, there is a less than 5% chance that a response will ever be created. Nevertheless, whenever possible, it is important that researchers use diligence and look for evidence that the dataset does not show evidence of an unusual distribution, especially one that is different from power-law. For example, the email responsiveness profile of an employee who has been away from email due to a three week holiday will not show a power-law distribution in the first few days after the holiday, and in that case the definition of online silence is not applicable.

Methodological Implications

A key factor in human communication research has been the ability to obtain large amounts of naturally occurring conversational records processed and ready for analysis. The work presented here highlights the potential that CMC holds for providing such data. We have shown that CMC conversations (“persistent conversation”) can be analyzed using tried and proven tools used for the analysis of face-to-face conversation and shares important attributes with traditional “mouth-to-ear” communication. Moreover, since the raw data of CMC is already digitized, and thus requires far less human effort to transform from the raw recordings to, for example, analyzable response times, far greater amounts of information can be processed, and results that are more robust quantitatively can be obtained. Moreover, datasets like those described in our work were collected unobtrusively (Webb, Campbell, Schwartz, & Sechrest, 2000), and thus represent a natural conversation. The availability of large datasets containing digitized and ready for analysis natural conversations, could revolutionize the methodology of studying human communication (Newhagen & Rafaeli, 1996).

Bibliography

Abbott, K., Mann, S., DeWitt, D., Sales, L., Kennedy, S., & Poropatich, R. (2002). Physician-to-physician consultation via electronic mail; The Walter Reed Army Medical Center Ask a Doc system. *Military Medicine, 167*(3), 200-204.

Adamic, L. A. (2005). *Zipf, Power-laws, and Pareto - a ranking tutorial*. Retrieved August 7, 2005, from http://www.hpl.hp.com/research/idl/papers/ranking/ranking.html

Aragon, S. R. (2003). Creating social presence in online environments. *New Directions for Adult and Continuing Education, 2003*(100), 57-68.

Axtell, R. L. (2001). Zipf Distribution of U.S. Firm Sizes. *Science, 293*(5536), 1818-1820.

Brady, P. T. (1965). The Technique for Investigating On-Off Patterns of Speech. *The Bell Systems Technical Journal, 44*(1), 1-22.

Brady, P. T. (1968). A Statistical Analysis of On-Off Patterns in 16 Conversations. *The Bell Systems Technical Journal, 47*, 73-91.

Caplan, S. E. (2003). Preference for online social interaction - A theory of problematic Internet use and psychosocial well-being. *Communication Research, 30*(6), 625-648.

Cappella, J. N. (1979). Talk-silence Sequences in Informal Conversations I. *Human Communication Research, 6*(1), 3-17.

Comellas, F., & Gago, S. (2005). A star-based model for the eigenvalue power law of internet graphs. *Physica A-Statistical Mechanics and Its Applications, 351*(2-4), 680-686.

Cramton, C. D. (2001). The Mutual Knowledge Problem and Its Consequences for Dispersed Collaboration. *Organization Science, 12*(3), 346-371.

Customer-Respect-Group. (2004). *The Customer Respect Group*. Retrieved August 27, 2004, 2004, from http://www.customerrespect.com

Danchak, M. M., Walther, J. B., & Swan, K. P. (2001). *Presence in Mediated Instruction: Bandwidth, Behavior, and Expectancy Violations.* Paper presented at the annual meeting on Asynchronous Learning Networks, Orlando, FL.

Davenport, T., & Beck, J. (2001). *The attention economy: understanding the new currency of business*. Boston: Harvard Business School Press.

Edelman, B. (2004, January 14, 2004). *Earnings and Ratings at Google Answers (preliminary draft)*, 2004, from http://cyber.law.harvard.edu/people/edelman/pubs/GoogleAnswers-011404.pdf

Erickson, T. (1999). Persistent Conversation: An Introduction. *Journal of Computer-Mediated Communication, 4*(4).

Erickson, T., & Herring, S. C. (2005). *Persistent Conversation: A Dialog Between Research and Design.* Paper presented at the 38th Hawaii International Conference on System Sciences, Big Island, Hawaii.

Erickson, T., & Kellogg, W. A. (2000). Social Translucence: An Approach to Designing Systems that Support Social Processes. *ACM Transactions on Computer-Human Interactions, 7*(1), 59-83.

Feldman, M. S., & March, J. G. (1981). Information in Organizations as Signal and Symbol. *Administrative Science Quarterly, 26*(2), 171.

Gabaix, X., Gopikrishanan, P., Pelrou, V., & Stanley, H. E. (2004). A theory of power-law distributions in financial market fluctuations. *Nature, 423*, 267 - 270.

Goldberg, D. S., Franklin, G., & Roth, F. P. (2005). *Breaking the Power Law: Improved Model Selection Reveals Increased Network Complexity.* Paper presented at the 13th Annual International conference on Intelligent Systems for Molecular Biology, Ann Arbor, MI.

Goodwin, C. (2002). Time in Action. *Current Anthropology, 43*(Supplement, August-October), S19-S35.

Hamilton, M. B. (2005). *Online Survey Response Rates and Times*. Lake Oswego, OR: Tercent, Inc.

Herring, S. C. (1999). Interactional coherence in CMC. *Journal of Computer-Mediated Communication, 4*(4).

Hirsh, L. (2002). *E-Tail Customer Service: Finally Working?* Retrieved May 30, 2004, from http://www.technewsworld.com/story/19353.html

Jaffe, J., & Feldstein, S. (1970). *Rhythms of Dialogue*. New York: Academic Press.

Jones, Q., Ravid, G., & Rafaeli, S. (2004). Information overload and the message dynamics of online interaction spaces: A theoretical model and empirical exploration. *Information Systems Research, 15*(2), 194-210.

Kalman, Y. M., & Rafaeli, S. (2005). *Email chronemics: unobtrusive profiling of response times.* Paper presented at the 38th Hawaii International Conference on System Sciences, Big Island, Hawaii.

Keeling, M., & Grenfell, B. (1999). Stochastic dynamics and a power law for measles variability. *Philosophical transactions of the royal society of London series b-biological sciences, 354*(1384), 769-776.

Lantz, A. (2003). Does the use of E-mail change over time? *International Journal of Human-Computer Interaction, 15*(3), 419-431.

Lewis, C. E., Thompson, L. F., Wuensch, K. L., Grossnickle, W. F., & Cope, J. G. (in press). The impact of recipient list size and priority signs on electronic helping behavior. *Computers in Human Behavior, In Press, Corrected Proof.*

Mattila, A. S., & Mount, D. J. (2003). The impact of selected customer characteristics and response time on E-complaint satisfaction and return intent. *International Journal of Hospitality Management, 22*(2), 135-145.

Matzler, K., Pechlaner, H., Abfalter, D., & Wolf, M. (2005). Determinants of response to customer e-mail enquiries to hotels: evidence from Austria. *Tourism Management, 26*(2), 249-259.

McLaughlin, M. L. (1984). *Conversation: How Talk is Organized*. Beverly Hills: Sage Publications.

McLaughlin, M. L., & Cody, M. J. (1982). Awkward Silences: Behavioral Antecedents and Consequences of the Conversational Lapse. *Human Communication Research, 8*(1), 299-316.

Mitzenmacher, M. (2003). A Brief History of Generative Models for Power Law and Lognormal Distributions. *Internet Mathematics, 1*(2), 226-251.

Morahan-Martin, J., & Schumacher, P. (2000). Incidence and correlates of pathological Internet use among college students. *Computers in Human Behavior, 16*(1), 13-29.

Negroponte, N. (1994). Talking with Computers. *Wired, 2*(3), http://www.wired.com/wired/archive/2.03/negroponte.html.

Newhagen, J. E., & Rafaeli, S. (1996). Why communication researchers should study the Internet: A dialogue. *Journal of Communication, 46*(1), 4-13.

Nonnecke, B., & Preece, J. (2000). *Lurker Demographics.* Paper presented at the Proceedings of CHI'2000, The Hague, Netherlands.

Pitkin, R. M., & Burmeister, L. F. (2002). Prodding tardy reviewers: a randomized comparison of telephone, fax, and e-mail. *JAMA, 287*(21), 2794-2795.

Poppel, E. (2004). Lost in Time: a Historical Frame, Elementary Procssing Units and the 3-second Window. *Acta Neurobiologiae Experimentalis, 64*, 295-301.

Qian, J., Luscombe, N., & Gerstein, M. (2001). Protein family and fold occurrence in genomes: Power-law behaviour and evolutionary model. *Journal of Molecular Biology, 313*(4), 673-681.

Raban, D. R., Ravid, G., & Rafaeli, S. (2005). *Paying for Answers: An Empirical Report on the Google Answers Information Market.* Paper presented at the AOIR: Internet Research 6.0: Internet Generations, Chicago, Il.

Rafaeli, S. (1984). *If the Computer is the Medium, What is the Message?* Paper presented at the Annual Conference of the International Communication Association, Honolulu, Hawaii.

Rafaeli, S. (1988). Interactivity: From new media to communication. In *Sage Annual Review of Communication Research: Advancing Communication Science* (Vol. Vol. 16, pp. 110-134). Beverly Hills, CA: Sage.

Rafaeli, S. (2004). In M. Consalvo, N. Baym, J. Hunsinger, K. B. Jensen, J. Logie, M. Murero & L. R. Shade (Eds.), *Constructs in the Storm*. New York, NY: Peter Lang.

Rafaeli, S., Ravid, G., & Soroka, V. (2004). *De-lurking in virtual communities: a social communication network approach to measuring the effects of social and cultural capital.* Paper presented at the 37th annual HICSS conference (Hawaii International Conference on System Sciences), Hawaii.

Ravid, G., & Rafaeli, S. (2004). A-synchronous Discussion Groups as Small World and Scale Free Networks. *First Monday, 9*(9).

Reed, W. J. (2001). The Pareto, Zipf and other power laws. *Economics Letters, 74*(1), 15-19.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1978). A Simplest Systematics for the Organization of Turn Taking for Conversation. In J. Schenkein (Ed.), *Studies in the Organization of Conversational Interaction* (pp. 7-55). New York: Academic Press.

Sheehan, K. B., & McMillan, S. J. (1999). Response Variation in E-Mail Surveys: An Exploration. *Journal of Advertising Research, 39*(4), 45-54.

Shenk, D. (1999). *The End of Patience: Cautionary Notes on the Information Revolution*. Bloomington, IN: Indiana University Press.

Stellin, S. (2003, June 30). Most Wanted: Drilling Down/Company Web Sites; Customer Care, Online. *The New York Times,* p. 8.

Strauss, J., & Hill, D. J. (2001). Consumer complaints by e-mail: An exploratory investigation of corporate responses and customer reactions. *Journal of Interactive Marketing, 15*(1), 63-73.

Tyler, J. R., & Tang, J. C. (2003). *When Can I Expect an Email Response? A Study of Rhythms in Email Usage*. Paper presented at the ECSCW 2003.

Vollarth, M., Kazenwadel, J., & Kruger, H. P. (1992). A Universal Constant in Temporal Segmentation of Human Speech. *Naturwissenschaften, 79*(10), 479-480.

Walther, J. B., & Tidwell, L. C. (1995). Nonverbal cues in computer-mediated communication, and the effect of chronemics on relational communication. *Journal of Organizational Computing, 5*, 355-378.

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (2000). *Unobtrusive Measures; Revised Edition*. Thousand Oaks, CA: Sage Publications Inc.

Zipf, G. K. (1949). *Human Behaviour and the Principle of Least-Effort*. Cambridge, MA: Addison-Wesley.