

L2 PRAGMATICS: EXPLORING LINGUISTIC, STYLISTIC AND PRAGMATIC FEATURES OF APOLOGIES AND REQUESTS IN CMC

Stephanie Michelet

Abstract

This paper uses discourse analysis to investigate the diverse pragmatic strategies that arise in English language learners' e-mail exchanges. Specifically, politeness, requests and apologies are analyzed. The data are taken from a series of six e-mail exchanges between native Quebecois speakers enrolled in an English language course as well as e-mails to their instructor. E-mail interaction is considered to lie somewhere between written and spoken discourse and therefore is a particularly interesting resource for understanding how language learners negotiate between levels of linguistic, stylistic and pragmatic appropriateness in producing their messages. The analyses show that language learners use a number of techniques to express politeness, make requests and apologize which demonstrate a range of perspectives regarding what is considered appropriate in time and (cyber) space.

Key words

pragmatics, speech acts, politeness, computer-mediated communication, E-mail, language learning

1 Introduction and research background

The purpose of this paper is to use discourse analysis to explore the range of strategies language learners use to express politeness, make requests and apologies in a computer-mediated communication environment. The data were collected from e-mail exchanges between language learners fulfilling a course assignment as well as spontaneous e-mails written to their professor. Considering the speech acts of requesting and apologizing, the author aims to identify the linguistic, stylistic and pragmatic features that occurred in the two types of exchanges.

Computer-mediated communication (CMC) is an increasingly common means for promoting authentic communicative tasks for language learners. Data collected from CMC has been used to describe various issues in language development including negotiation of meaning (Fernandez-Garcia & Martinez-Arbelaiz 2002, Lee 2001), oral performance (Abrams 2003, Payne & Whitney 2002) and grammatical competence (Pelletieri 2000). Relatively few studies have

considered how internet communication can be used to understand (Biesenbach-Lucas 2007, Chen 2006) or promote pragmatic development (Belz & Vyatkina 2003, 2005). However, given the prevalence of CMC and its capacity to facilitate communication between people of different cultural and linguistic backgrounds, such pragmatic issues are not only interesting, but also relevant to real life language use. Therefore it is increasingly evident that researchers and educators need to “understand the norms of language use during computer-mediated interaction and their potential relationship to second language acquisition (SLA)” (Smith 2003: 38).

Although there have been a few studies that gather naturally-occurring data (Beebe & Cummings 1996, Bou-Franch & Lorenzo-Dus 2005) most studies regarding NNS requests (Blum-Kulka 1991, Blum-Kulka, House & Kasper 1989) and apologies (Cohen & Olshtain 1993) use data from discourse completion tasks. This paper contributes to the growing but limited amount of research that uses computer-mediated communication as a source of authentic, naturally occurring speech acts (Biesenbach-Lucas 2007, Chen 2006).

2 Internet features

Computer-mediated communication (CMC) combines features and functions that are characteristic of written and oral language (Chun 2003, Davis & Brewer 1997, Shekary & Tahririan 2006, Vinagre 2005), and also has specific features of its own. As a written form, CMC lacks intonation, produces a permanent record of discourse, can be lexically dense, and relies on punctuation and textual formatting. But CMC also has features that imitate oral discourse such as the use of italics and bold face type to indicate stress on words and phrases, first person perspective, and a tendency towards informality. Finally, CMC has a particular style of its own as evidenced by the acceptance of surface errors, the use of simplified registers and syntax, which results in shorter sentences, as well as abbreviations and other reduced forms of words, and the use of emoticons and symbols, such as smiley faces to indicate emotions and gestures (Chen 2006, Lee 2001, Smith 2003). These features pose challenges to e-mail users who must choose among conflicting notions of appropriateness in internet communication which can appear, on the one hand, to be free from the strict rules of formal writing, yet intimidating in its lack of standard rules of what is appropriate, on the other (Chen 2006).

The acquisition of pragmatics requires knowledge of linguistic and social rules, and we are just beginning to understand how learners negotiate among these different skills in the CMC context. Towards this goal this paper identifies

and describes the types of strategies that a group of English-as-a-second-language (ESL) learners used to make apologies and requests and express politeness in a computer-mediated communication environment.

3 Project description

The participants were native French speakers enrolled in two sections of an ESL class. The courses met on different nights and were taught by the same teacher. The students' ages ranged from 20 to 60 years with an average age of 36. Their proficiency levels ranged from low to mid-intermediate, with various language learning experiences. For example, some students had traveled, or even lived and worked in English speaking contexts, while others had never left Québec. The data were drawn from two sources: and e-mail exchange project between the students, and e-mails written to the professor.

E-mails written to the professor were saved in a file and later transferred to a Word document. For the e-mail project, students were assigned partners from the other class and were matched for similar interests, proficiency level and gender as much as possible. The project consisted of the exchange of six e-mails over a six-week period. The topics were provided and reflected grammar and language functions (e.g. Describe a movie, show, performance, play, dance spectacular, art display, short trip, sports event, or wonderful meal that you recently enjoyed. Describe the experience using past tense verb forms, and then try to persuade your partner to do this activity. Give advice on what to do or not do regarding this activity). Students were asked to correct their partner's e-mails paying particular attention to the verb tenses specified in the assignment, send the corrections to their partner and print a copy to turn in to the instructor.

The assignments did not require requests or apologies; so all of the data collected here are derived from the students' own initiative to include either in their e-mail. Reasons for apologizing between students included not sending an e-mail on time, or not doing an assignment correctly. Apologies to the professor ranged from forgetting to include an attachment to missing a class. Student-student requests were limited; there were requests to have an e-mail sent to a particular address or on a certain day. The majority of the requests were written to the professor. These included everything from printing advice to grade revisions.

The data are comprised of 200 e-mails between students, based on the consent of the learners. They were printed and collected each week over a six-week period. E-mails to the professor were collected over the entire fourteen-week semester. Due to the exploratory nature of this study and its reliance on naturally occurring examples the body of the data is limited. In total 16 apologies and 30 requests were produced.

4 Data analysis

In this paper, politeness is analyzed according to Brown and Levinson's (1987) concept of face, or one's self-image. Politeness is a measure of awareness and respect for another person's negative face, which desires independence and does not want to be imposed upon, and positive face, which involves the need for acceptance and feelings of group membership. Acting appropriately in terms of face requires understanding how direct or indirect to be in accordance to one's negative and positive face.

Apologies and requests were coded according to the Cross-Cultural Speech Act Realization Project's (CCSARP) coding manual (Blum-Kulka et al. 1989). Blum-Kulka et al. identify seven apology strategies (cited in Kasper & Rose 1999: 86):

1. Offering an illocutionary force indicator device (IFID)
2. Assuming responsibility for the offence
3. Downgrading responsibility, or severity of offence
4. Offering repair
5. Expressing concern for the offended party
6. Appeasing the offended party
7. Promising forbearance

Politeness in requests was considered a function of two elements: directness and modality markers. Requests were coded according to the CCSARP (Blum-Kulka et al. 1989) as being direct (*Please, write me back*), conventionally indirect (*Can you write me back*), or non-conventionally indirect (*Do you have my e-mail address*). Specific strategies within these categories will be discussed as needed. Modality markers affect the level of politeness. Examples of modality markers include downgraders, or devices used to minimize the impact of the utterance on the hearer, and upgraders, which increase the force of the impact (House & Kasper 1981: 166-170).

5 Results and discussion

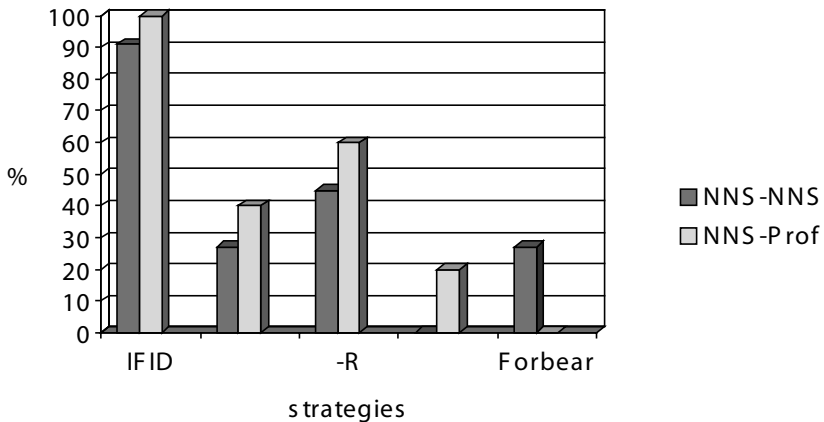
5.1 Discourse style

Irrespective of interlocutor, students tended to use a conversational style, characterized by simplified or reduced forms (*I just wanna know ...*, *Do you mind if I gonna ...*), contractions (*I'll*, *won't*), subject "I" deletion (*So sorry*), and paralinguistic symbolization including uppercase (*I didn't REWRITE her paper*),

multiple punctuation marks (“...”, “???”, “!!!”), and emoticons (such smiley face ;-)). Examples will be presented and discussed as relevant.

5.2 Apologies

In apology situations, there had been a violation of a social norm which led to “an action or utterance which is intended to ‘set things right’ ” (Trosborg 1987: 148). The nature of the students’ assignment in this study required the timely exchange of e-mails, a norm, which to the distress of the students, was often violated. Graph 1 shows the distribution of apology strategies found in this study, percentages were determined by dividing the number of strategies used by the total number of apology speech acts.



Graph 1: Percentages of apology strategies.

IFID = Illocutionary Force Indicator Device, +R = assuming responsibility, -R = downgrading responsibility, Repair = offering repair, Forbear = offering forbearance.

Eleven examples of apologies between students were collected, representing strategies 1, 2, 3, and 7, or some combinations of these. So while there were eleven total apologies between students the strategies appeared a total of 21 times. Ten of the eleven apologies included an illocutionary force indicator device (IFID). The second most common strategy was an IFID followed by strategy 3 or downgrading responsibility by offering an explanation or account. Example (1) uses an IFID and assumes responsibility.

(1) *S’cuse my late for this e-mail!*

Despite using two strategies, certain characteristics of the style, the informality

and brevity of the message, as indicated by the reduced form of the word *excuse*, and the lack of an explanation create a less polite effect.

Four of the eleven apologies used an IFID followed by downgrading responsibility. Trosborg (1987) refers to this as an explanation or account. In this strategy “a complaineer admits that what he has done was undesirable, but he tries to lessen the blame which can be attached to him by referring to mitigating circumstances that may excuse his behaviour. Thus an explanation or an account serves as an ‘excuse’ for a committed offence ...” (Trosborg 1987: 151). This strategy is demonstrated in examples (2) and (3).

(2) *I'm very sorry for delay. I'm improve a lot of problem with my own computer. I had to change it. Now every thin[g] i[s] under control.*

(3) *Excuse me for my last email. I didn't understand my homework, I was sick.*

Example (2) uses the intensifier *very* along with an explanation that implies the student actively worked to overcome the problem, and provides an indirect promise of forbearance in the statement “Now everything is under control”, which implies that there will not be any more late e-mails. In contrast, example (3) gives two excuses: not understanding the homework and being sick; however; this seems less effective because the apologizer assumes no responsibility for finding out how to do the assignment correctly. Missing class due to illness or unknown reasons was offered as an excuse three times.

Finally (4) and (5) are two of three examples of promising forbearance. This strategy involves promising not to repeat the offence or offering ways of improving one's behavior (Trosborg 1987):

(4) **You will always receive my e-mail on Wednesday because I have too much work to do at school with my students. Sorry!** (emphasis original)

(5) *I will try to send you my email before Tuesday, but sometimes I don't have time. I have other courses on Monday and these last weekends I wasn't at home. Maybe it will be ok on next weeks.*

In (4) we see the use of bold-face type and underlining along with an exclamation mark after the IFID. These are clearly meant to intensify the speech act. In contrast, (5) lacks an IFID and uses language such as *will try*, *sometimes*, and *maybe* that minimize the force of the speech act. The lack of an IFID in (5) emphasizes the lack of responsibility provided in the explanations, which is also implied by the lack of commitment in respect to the promise of forbearance.

A total of five apologies were offered to the professor including strategies 1, 2, 3, and 4. In contrast to the student-student e-mails, apologies to the professor included more emphasis via exclamation marks as in examples (6-8):

(6) *Sorry I forget to send you a copy!!!*

(7) *I'm sorry!*

(8) *So sorry and see you tonight !*

In addition, students used the complete phrase “I’m/I am sorry” three times with the professor but never to each other. Students also tended to give lengthier explanations to the professor as in (9):

(9) *I am sorry about yesterday; i couldn't go to English course.*

My little son was sick; he was doing too much fever.

I had to look after him all night. Fortunately he is better today.

Apologies tended to be brief perhaps as a way to minimize responsibility and save face. Students relied on the IFID as the main apology strategy; although, other strategies were added to support the IFID especially with the professor.

5.3 Requests

Requests can be characterized as a “pre-event” which comes at some cost to the hearer (House & Kasper 1981). Acknowledging this cost via level of directness and the use of modality markers is a function of politeness. The use of modality markers was limited in scope and number. Students relied primarily on the politeness marker *please*. In order to downgrade requests, there was some use of an agent avoider (*Would it be possible*) and understaters (*I have a little problem; I just ...*). Upgraders were even more rare, except for the case of time intensifiers (*as soon as possible*). Other upgraders included: intensifier (*probably*) and plus committer (*I'm sure*). In both data sets, the most common conventionally indirect strategy is the query preparatory which queries a preparatory condition of a requested action (House & Kasper 1981). For example, the preparatory conditions could be the hearer’s ability or willingness to fulfill a request (Blum-Kulka et al. 1989). A discussion of the students’ strategies to this effect follows.

Requests among students were straightforward and unelaborated and lacked politeness markers, as in (10) and (11):

(10) *You can use my e-mail at my firm. It's easier for me to print it.*

(11) *Could you send your email at (address).*

Example (10) uses a state preparatory strategy, which opposed to a query preparatory, asserts a preparatory condition of a requested action (House & Kasper 1981). However, one student diverged from the predominately informal and conversational style of CMC and produced a request in accordance to linguistic and stylistic rules of formal writing:

(12) *I think that it's difficult for me to correct and print your e-mail before my course if you send me your e-mail the thursday. Could you please send me your e-mail before the wednesday 6H00 pm because, the day*

after I work and I go to school the evening. Thank you very much for your understanding!

P.S. I know that you're very busy! ;)

In (12) a grounder including a conditional clause sets the stage for the request which is conventionally indirect, including the politeness marker *please* and an additional grounder. The student also expresses gratitude, followed by a disarmer. The winking, smiley-face is an indicator of solidarity via positive politeness, suggesting that despite the seriousness and formality of the request the student wishes to maintain a friendly relationship. Clearly this student spent a lot of time composing this request with an interest in showing respect for the partner while being clear about the reasons for the request. The effectiveness of this strategy is evidenced by the response, see example (4) above.

Requests to the professor included direct, conventionally indirect, and nonconventionally indirect strategies combined with various supporting moves. Due to space restrictions only direct and conventionally indirect requests are discussed. While there was one example of an imperative (*Give me some news about that please*), the majority of direct requests to the professor were want statements. Want statements are considered more direct and thus less polite than other request strategies; however, students modified their requests in ways that increased their politeness. Want statements “state the speaker’s desire that the hearer carries out the act” (Blum-Kulka et al. 1989: 18).

(13) *I would like to have my finale note for this course if you want! Can you gave me please?*

In (13) the student’s use of a direct request in the form of a want statement is immediately mitigated by an imposition minimizer (term from Blum-Kulka et al. 1989) that explicitly queries the professor’s willingness (*if you want!*). The latter strategy was observed to be very frequent among French Canadians in a cross-cultural study by Blum-Kulka (1989), and should be considered an example of transfer from the French verb *vouloir*. This also occurred as (*Would you like to ... e. g. Would you like to send me the homeworks for this week please?*). The initial request is further mitigated by a conventionally indirect query preparatory strategy with a politeness marker. Other ways that want statements were modified included grounders (14), and downtoners (15).

(14) *Tonight, I won't be in the class, because I'll have a christmas party of my job. I would like to know content for the final exam.*

(15) *I just wanna know if we have an oral presentation Thursday ... Many people called me to know if we have to prepare it or not?*

Several examples show a complex interaction among strategies, modifiers and perspective as in (16) and (17).

- (16) *I would like to know if that would be possible that I make presentation alone with you, because I am really embarrassed and I am very shy. Can you send me your answers as soon as possible.*

In (16) the student uses a conditional clause, an agent avoider and a grounder to increase the politeness of the direct request. This is followed by a conventionally indirect request followed by a time intensifier, which is less polite as it puts pressure on the professor to respond, thereby threatening the professor's negative face.

- (17) *I just wanna know if I have to correct the letter and write the past progressive and some clauses because (partner's name) wrote all her letter and used only the past tense ... She didn't put some clauses (when, while as soon as) I corrected the structures and vocabulary as the better I can but I didn't REWRITE her letter. I hope it's not what I have to do????*

In (17) the student uses several features of computer-mediated communication: a familiar, conversational style, reduced forms (*wanna, it's*), and paralinguistic symbols ("...", "????", REWRITE). The primary direct want statement request is minimized with an understater (*just*), and finally the request is reworded into another direct request (*I hope it's not what I have to do*). The four question marks have an interesting impact on the request because the structure of the request is not a question. It is a statement of expectation; however, the use of question marks indicates uncertainty, giving the professor the option to reply that yes indeed that is what has to be done. Compare, for example, the impact of replacing the question marks by exclamation marks: *I hope it's not what I have to do!!!!* Meanwhile the repetition of the question marks does add intensity to the request, which simultaneously puts pressure on the professor to fulfill the student's expectation.

Conventionally indirect requests are considered more polite because they show more distance and therefore respect for negative face. The students' main strategies for conventionally indirect requests were query preparatory. As described above, query preparatory requests refer to the ability, willingness and possibility of the interlocutor to fulfill the request (Blum-Kulka et al. 1989). In all of the indirect requests except for two, the choice of perspective is hearer-oriented. Examples (18-20) represent the range of query preparatory requesting strategies that students used.

- (18) *I have a little problem.
I don't know how to print my contact e-mail.
Please could you let me know?*
- (19) *I'm sure some people will be absent next oct 31, so it will be possible*

to do a short review about phrasal verbs the course after?

(20) *Do you mind if I gonna miss 4 english class.*

Example (18) is among the most typical types of request, including a preparator and grounder. In (19) an upgrader is used (*I'm sure*) followed by the impersonal which softens the impact of the intensifier. Finally, in (20) there are reduced forms (*gonna* and *4*), copula deletion, and a lack of punctuation. This request was embedded in a long e-mail about a job opportunity abroad written in a very informal way.

Students' requests show an eclectic mix of written, oral and CMC style features, often all in the same speech act. Negative politeness is affected by the lack of downgraders and use of direct requests; however, the use of positive politeness, which is more characteristic of CMC, is prevalent, and at times appropriate.

6 Conclusion

This investigation set out to analyze English language learners' politeness, request and apology strategies in terms of their linguistic, stylistic and pragmatic features in a computer-mediated communication context. The data were drawn from naturally occurring language produced in student-student (S-S) as well as student-professor (S-P) e-mails. According to the data, it seems that CMC affects strategy selection and may explain a tendency toward multiple strategy use. Examples of multiple strategy use found in this study include the integration of oral, written and internet styles, different linguistic features such as modality markers, and fluctuating positive and negative politeness features. It was observed that S-S requests and apologies tended to be more direct and lacking politeness markers. The S-P e-mails encompassed a wider range of strategies, and the expression of both positive and negative politeness.

In addition to the CMC context, several additional factors must be taken into account to explain the observed differences. First, the data were drawn from two fundamentally different tasks. The S-S e-mails were written as part of a course assignment that required a certain level of pre-planning and focus on form as they would eventually be read and evaluated by the professor. On the other hand S-P e-mails were written spontaneously and driven by a need originating from the student him- or herself. This means that different levels of monitoring could have been taking place that affected the way the students approached each type of task. In this respect, one would expect differences in the language produced, and the data indeed bear this out. Future research should consider how the nature of the task interacts with strategy use in CMC.

Other influences on the students' strategies include first language norms for politeness such as the tendency among the Quebecois to use query preparatory requests. Proficiency level is also indicated by examples of formulaic strategies directly translated from the L1 to the L2. Further study is needed to understand the influence of proficiency, culture and linguistic background on L2 speech acts in CMC.

This exploratory study provided evidence that the CMC environment influences politeness, and the apology and request strategies of language learners. Studying internet communication is a promising means for analyzing the pragmatic strategies of language learners in an authentic linguistic context. Future research should concentrate on collecting more data in comparable situations, which would permit a comprehensive investigation of strategy use. Such studies would contribute to our understanding of pragmatic development and aid in the development of pedagogical materials.

References

- Abrams, Z. (2003) 'The effect of synchronous and asynchronous CMC on oral performance in German.' *The Modern Language Journal* 87/2, 157-167.
- Bardovi-Harlig, K., Dörnyei, Z. (1998) 'Do language learners recognize pragmatic violations? Pragmatic versus grammatical awareness in instructed L2 learning.' *TESOL Quarterly* 32/2, 233-262.
- Beebe, L., Cummings, M. (1996) 'Natural speech act data versus written questionnaire data: How data collection method affects speech act performance.' In: Gass, S., Neu, J. (eds) *Speech Acts across Cultures*. New York: Mouton de Gruyter. 65-86.
- Belz, J., Kinginger, C. (2003) 'Discourse options and the development of pragmatic competence by classroom learners of German: The case of address forms.' *Language Learning* 53/4, 591-647.
- Belz, J., Vyatkina, N. (2005) 'Learner corpus analysis and the development of L2 pragmatic competence in networked intercultural language study: The case of German modal particles.' *The Canadian Modern Language Review/La Revue canadienne des langues vivantes* 62/1, 17-48.
- Biesenbach-Lucas, S. (2007) 'Students writing emails to faculty: An examination of e-politeness among native and non-native speakers of English.' *Language Learning & Technology* 11/2, 59-81.
- Blum-Kulka, S. (1989) 'Playing it safe: The role of conventionality in indirectness.' In: Blum-Kulka, S., House, J., Kasper, G. (eds) *Cross-Cultural Pragmatics: Requests and Apologies*. Norwood, NJ: Ablex Publishing Corporation. 37-70.
- Blum-Kulka, S. (1991) 'Interlanguage pragmatics: The case of requests.' In: Phillipson, R., Kellerman, E., Selinker L., Sharwood-Smith, M., Swain, M. (eds) *Foreign/Second Language Pedagogy Research*. Clevedon, UK: Multilingual Matters. 255-272.
- Blum-Kulka, S., House, J., Kasper, G. (1989) 'The CCSARP Coding Manual.' In: Blum-Kulka, S., House, J., Kasper, G. (eds) *Cross-Cultural Pragmatics: Requests and Apologies*. Norwood, NJ: Ablex Publishing Corporation. 273-294.

- Brown, P., Levinson, S. (1987) *Politeness: Some Universals in Language Use*. Cambridge: Cambridge University Press.
- Bou-Franch, P., Lorenzo-Dus, N. (2005) 'Enough DCTs: Moving methodological debate in pragmatics into the 21st century.' Paper presented at the *16th International Conference on Pragmatics and Language Learning*. Bloomington, IN (manuscript).
- Chen, C.-F. E. (2006) 'The development of e-mail literacy: From writing to peers to writing to authority figures.' *Language Learning & Technology* 10/2, 35-55.
- Chun, L. (2003) *Text-based online chat, noticing and second language acquisition*. Research Proposal. Online document: <http://www.msu.edu/user/laichun1/researchproject.htm>
- Cohen, A., Olshtain, E. (1993) 'The production of speech acts by EFL learners.' *TESO Quarterly* 27/1, 33-56.
- Cohen, A., Shively, R. (2007) 'Acquisition of requests and apologies in Spanish and French: Impact of study abroad and strategy-building intervention.' *The Modern Language Journal* 91, 189-212.
- Davis, B., Brewer, J. (1997) *Electronic Discourse: Linguistic Individuals in Virtual Space*. Albany, NY: SUNY Press.
- Fernandez-Garcia, M., Martinez-Arbelaz, A. (2002) 'Negotiation of meaning in nonnative speaker- nonnative speaker synchronous discussions.' *CALICO Journal* 19/2, 279-294.
- House, J., Kasper, G. (1981) 'Politeness markers in English and German.' In: Coulmas, F. (ed.) *Conversational Routine: Explorations in Standardized Communication Situations and Prepatterned Speech*. The Hague: Mouton Publishers. 157-185.
- Kasper, G., Rose, K. (1999) 'Pragmatics and SLA.' *Annual Review of Applied Linguistics* 19, 81-104.
- Lee, L. (2001) 'Online interaction: Negotiation of meaning and strategies used among learners of Spanish.' *ReCALL* 13/2, 232-244.
- Payne, S., Whitney, P. (2002) 'Developing L2 oral proficiency through synchronous CMC: Output, working memory, and interlanguage development.' *CALICO Journal* 20/1, 7-32.
- Pelletieri, J. (2000) 'Negotiation in cyberspace: The role of chatting in the development of grammatical competence.' In: Warschauer, M., Kern, R. (eds) *Network-Based Language Teaching: Concepts and Practice*. Cambridge: Cambridge University Press. 59-86.
- Rose, K. (2000) 'An exploratory cross-sectional study of interlanguage pragmatic development.' *Studies in Second Language Acquisition* 22, 27-67.
- Shekary, M., Tahririan, M. (2006) 'Negotiation of meaning and noticing in text-based online chat.' *The Modern Language Journal* 90/4, 557-573.
- Smith, B. (2003) 'Computer-mediated negotiated interaction: An expanded model.' *The Modern Language Journal* 87/1, 38-57.
- Sotillo, S. (2005) 'Corrective feedback via Instant Messenger learning activities in NS-NNS and NNS-NNS dyads.' *CALICO Journal* 22/3, 467-496.
- Trosborg, A. (1987) 'Apology strategies in natives/non-natives.' *Journal of Pragmatics* 11, 147-167.
- Vinagre, M. (2005) 'Fostering language learning via email: An English-Spanish exchange.' *Computer Assisted Language Learning* 18/5, 369-388.

ON NARROW P-THEME PARAGRAPH IN FICTION, JOURNALISTIC AND ACADEMIC TEXTS

Renata Pípalová

Abstract

The paper deals with the thematic organization in paragraphs. It elaborates on the paragraph typology pioneered by Mathesius (1942) and Daneš (1994, 1995), and draws a distinction between two paragraph supratypes, viz. the narrow P-theme paragraphs and their broad P-theme counterparts. Focussing on the former, the study explores their register-specific tendencies detected in a corpus of authentic British English texts.

Key words

paragraph, narrow P-theme paragraph, broad P-theme paragraph, thematic build-up, register, corpus, subcorpus, central, non-central, stable P-theme paragraph, unfolded P-theme paragraph

0 Introduction

This paper explores the thematic build-up of paragraphs. It follows mainly the relevant Czech linguistic tradition.

In his pioneering study, Mathesius (1942/1982: 144) maintains that the integrity of the paragraph follows from the existence of the Paragraph Theme. He distinguishes between three types of paragraphs, viz. paragraphs where the Paragraph Theme is stable, where it unfolds, and where it develops.

Mathesius's original typology is crucially elaborated on by Daneš (1994/1995) who investigates particularly the interrelation between the Paragraph Theme (P-theme) and the utterance themes (U-themes). He explores the ways the individual U-themes contribute to the construction of the P-theme. Employing his own concept of thematic progressions (see e.g. Daneš 1974), he distinguishes in all four types of paragraphs differing in their thematic build-up. His typology includes paragraphs with a stable P-theme, paragraphs the P-themes of which are unfolded by a number of particular U-themes, paragraphs with a content frame and paragraphs in which the P-theme develops. It should be noted that Daneš examines solely academic paragraphs.

1 The present approach

Somewhat elaborating on Daneš's approach, Pípalová (2008) seeks to establish the above paragraph typology on a wider range of tentative criteria. In

her framework, the major paragraph types are conceived of as aggregates of the characteristics produced by various mutually interlocked criteria.

“If the cluster of the proposed criteria is comprehensive enough, it includes at least the relationship between the P-theme and the (U-/FSP) themes¹, (i.e. the degree of in/stability of choices from the Thematic area², affecting the range of the thematic Discourse Subjects³ and their variability across the paragraph); the types, arrangement and hierarchy of thematic progressions⁴; the type of thematic paradigm (the cohesive ties⁵ and means involved in thematic units); the role assigned to thematic progressions and cohesive links (their incidence); and the placement and characteristics of the P-theme exponents. Naturally, additional criteria may be revealed once paragraphs of various build-up patterns are carefully scrutinized and contrasted with one another.

The above criteria are seen as working in concert in the apperception of paragraph typology. Some paragraphs may be delimited only on the basis of a single criterion; others, however, may be founded on much firmer grounds.

Furthermore, the original paragraph types may be arranged along a relatively continuous cline. The two opposite ends of this cline differ radically in their build-ups and epitomize two distinct configurations of features. Thus, in this paper we propose to distinguish between two crucial supratypes, viz. Narrow and Broad P-theme paragraphs.”

(Pípalová 2008: 38)

In other words, when the above criteria are applied to a larger corpus, in view of their build-up, paragraphs will form a relatively smooth scale. On this scale, two conspicuous centres (cores), corresponding to the two paragraph supratypes, may be identified, differing remarkably in the cluster of mutually interlocked features. It goes without saying that between these two centres (cores), there is a boundary and that each core alone displays its own centre and its own periphery.

On this occasion, attention will be given solely to a single supratype, viz. the narrow P-theme paragraphs. Section 2 provides their succinct theoretical description and illustration. Section 3 discusses their distribution in a corpus of British English texts.

2 Narrow P-theme paragraphs

2.1 Theoretical description

The narrow P-theme paragraphs are those embodying the top degrees of stability in their thematic build-up. Hence they are rather focussed, foregrounding⁶ a single major Thematic Discourse subject (DS). In other words, the content of their P-theme is considerably hierarchized and due to that, the authors focus on its narrowest layer, i.e. on what they choose to see as its prominent, salient DS, i.e. a dominant Thematic entity.

The explicit exponent of the foregrounded Thematic DS is typically launched through the Topic Sentence which usually appears paragraph-initially. It motivates the U-thematic functions in all the paragraph utterances. Therefore, a clearly prevalent arrangement is paradigmatic.

It follows that the foregrounding of the above-mentioned dominant Thematic DS is achieved primarily by the extreme stability of choices from the Thematic area. Such constrained selection from the Thematic area is matched by a considerable degree of thematic as well as cohesive bonding. Indeed, the Thematic DS exponents are interlaced by thematic progressions and cohesive links. Due to their consistent use, all paragraph utterances are bonded through their prominent thematic elements.

Moreover, there is a pronounced constraint on the range of such bonding. As a rule, to foreground the prominent Thematic DS, authors employ a limited variety of both, thematic strings and cohesive links, interrelating such DSs. Naturally, the U-themes tend to be interwoven by paradigmatic progressions, and the cohesive links between them establish, and never exceed, the narrow thematic cohesive paradigm (embracing identity, equivalence, and inclusion links). As a rule, the introduction of thematic DSs is linguistically motivated. On the recoverability scale (Geluykens 1991), the thematic elements rank among those which are either recoverable directly or almost directly.

Generally speaking, the narrow P-theme paragraphs tend to be perceived as inducing coherence rather strongly. (Their nearest relatives would be content frame paragraphs with a dominant Thematic DS. These are also frequently constructed paradigmatically and do feature a salient Thematic DS. However, the cohesive links establish the broad paradigm and the paragraphs thematize also components from the broader layers of the Theme.)

In principle, authors of narrow P-theme paragraphs may employ two distinct, though by no means unrelated methods of foregrounding the dominant Thematic DS. Either its exponent is simply reiterated throughout the U-themes (Examples I-V below) or the individual U-themes are related to the dominant Thematic entity as its parts, constituents, examples, subtypes, species, instances, etc. (Examples

VI-IX below). The former subset complies with the stable P-theme paragraphs, whereas the latter with paragraphs where the P-theme is unfolded. Naturally, the foregrounding is more pronounced in the stable P-theme pattern than in the unfolded P-theme pattern. (For more on the narrow P-theme paragraphs, see Pípalová 2008 and Pípalová forthc.).

2.2 Methodology and illustration

The paragraphs in question are illustrated below. It should be stressed that examples I-IX are central and come from all the register corpora, i.e., academic (A), journalistic (J) and fiction (F) ones (see also 3.0 below). For the sake of simplicity, in this paper the unit of analysis is the main (independent) clause. Whenever there are more thematic units per a main clause, priority is naturally given to that/those which is/are more directly relevant to the P-theme⁷. Such unit(s) tend(s) to be suggested by bonding. Moreover, sometimes it is possible to distinguish between more local links and those critical for the overall integrity of the paragraph. Although this article focuses solely on the thematic build-up of the paragraphs in question, it should be pointed out that coherence can also exploit various other devices (e.g., explicit connectives, syntactic parallelism, etc.), and be further enhanced by some such means. However, as is well known, the perception of coherence need not be established on the existence of any such links and devices at all.

- (I) 1. *Birmingham, for so long the ugly duckling of British cities, has become the country's third most popular destination for foreign visitors.* 2. *The city that was once synonymous with mugger-infested subterranean walkways and impenetrable urban motorways is now more popular with overseas tourists than anywhere except London and Edinburgh.* (J2, 57)
- (II) 1. *A few hundred years ago, smallpox killed as many people as cancer or heart disease do today.* 2. *It was finally wiped out by an immunisation programme using a close but less virulent relative of smallpox, called vaccinia.* (J1, 78)
- (III) 1. *In all her life Cinderella had never known such happiness as she danced.* 2. *But not for a moment did she forget her fairy godmother's warning.* 3. *And when the big palace clock chimed a quarter-to-twelve, she slipped out of her Prince's arms and fled.* (F1, 44)
- (IV) 1. *Harry couldn't think of anyone who deserved to win a large pile of gold more than the Weasleys, who were very nice and extremely poor.* 2. *He picked up Ron's letter and unfolded it.* (F2, 43)
- (V) 1. *Rivers, lakes and other wet habitats have a special place in Britain's*

- natural heritage. 2. Apart from their botanical importance, they provide some of the most vulnerable sites for bird life. 3. They have in many cases escaped the changes due to forest cover and (until recently) those brought about by agricultural improvement. (A1, 32)*
- (VI) *1. Tomasz Gliszczynski and Wacław Zawadowski, statistics teachers at the Academia Podlaska in Siedlce, received Belgian Euro coins from Poles returning from jobs in Belgium and immediately set their students spinning them. 2. Gliszczynski says spinning is a more sensitive way of revealing if a coin is weighted than the more usual method of tossing in the air. (J4, 12)*
- (VII) *1. Two years ago Emma was diagnosed with osteosarcoma, a rare form of bone cancer, after doctors found a 12in tumour in one of her legs. 2. The family had never heard of the disease which, they were told, affects 30 young people in the UK every year. (J1, 52)*
- (VIII) *1. The porcupine (bold in the original, R.P.) is nocturnal and capable of doing much harm to our native flora and fauna. 2. A number of Himalayan porcupines escaped in 1969 from a wildlife park near Okehampton in Devon, and although there are now probably less than 20 in the wild, they still seriously damage the local conifer plantations. 3. In 1972 a pair of North African crested porcupines escaped from Alton Towers in Staffordshire. 4. They may have bred, but apparently failed to establish a colony. (A1, 115)*
- (IX) *1. Late Norman architecture is indeed a personal and in some ways rather 'Baroque' style. 2. At St Frideswide's Priory, Oxford (now Christ Church Cathedral), the arcade is based on the giant columns of the Tewkesbury type; the triforium is not placed above them, however, as one should expect, but is oddly tucked between them. 3. At another Augustinian house, Worksop (or Radford) Priory, Nottinghamshire, there is a similarly complex and unexpected internal elevation in which the tribune has a busy tripartite rhythm with a huge central arch cutting into the clerestory area. 4. In the nave walls of St David's Cathedral, Dyfed, begun in 1180, a remarkable spacial effect is achieved. 5. Two-storey window arches contain within them both the triforium and its passage, and a large clerestory passage. (A2, 21)*
- (X) *1. Harry, though still rather small and skinny for his age, had grown a few inches over the last year. 2. His jet-black hair, however, was just as it always had been: stubbornly untidy, whatever he did to it. 3. The eyes behind his glasses were bright green, and on his forehead, clearly visible through his hair, was a thin scar, shaped like a bolt of lightning. (F2, 28)*

3 Research

3.0 Corpus description

This paper is based on the results of a larger-scale study exploring paragraphs in authentic texts. The overall research corpus features 2,070 graphic paragraphs of contemporary British English. It consists of three corpora which correspond to three distinct functional styles/broad stylistic varieties/registers⁸, namely academic writing, journalism and fiction. Each register corpus comprises 690 paragraphs and is made up of two subcorpora, each exhibiting 345 paragraphs. For the sake of clarity, the academic corpus includes a subcorpus of samples dealing with the subject matter of humanities and another dealing with natural sciences; the journalistic corpus comprises newspaper and magazine subcorpora, and the fiction corpus is composed of children's and adult literature subcorpora. In order to reduce somewhat the impact of the authors' idiolects, every register subcorpus includes three distinct source subcorpora, each encompassing 115 continuous graphic paragraphs.

3.1 Narrow P-theme paragraphs in the corpus

Of the research corpus investigated, the narrow P-theme paragraphs were found to form a somewhat larger part. More specifically, they amounted to 1,108 paragraphs, i.e. 53.53 per cent of the overall corpus.

3.2 Narrow P-theme paragraphs in individual register corpora and subcorpora

Of the corpora surveyed, the top representation of the narrow P-theme paragraphs was attested in fiction. In fact 476 instances were collected, corresponding to 68.99 per cent of the fiction corpus. The share of the narrow P-theme paragraphs was somewhat lower in journalism, slightly exceeding a half of the specimens in question. More specifically, there were 391 instances in all, i.e. 56.67 per cent of the journalism corpus. Finally, the incidence of the narrow P-theme paragraphs was found to be the lowest of all in academic samples. There were only 341 such paragraphs, corresponding to 34.93 per cent of the academic corpus.

Scrutinizing now the individual subcorpora, the researcher cannot but notice striking disproportions. The top share of the narrow P-theme paragraphs was detected in the newspaper subcorpus. There were altogether 287 instances,

corresponding to 25.90 per cent of all. The second most prolific subcorpus turned out to be that of adult fiction, from which 256 instances were drawn (23.10% of all). With its 220 instances (19.86% altogether), children's fiction was found to be also a relatively ample source of the narrow P-theme paragraphs. Next came the academic subcorpus of natural sciences, exhibiting 144 cases (12.99% altogether). It was followed by the magazine subcorpus (104 paragraphs, corresponding to 9.39% in all). The narrow P-theme paragraphs were the rarest of all in the academic subcorpus of humanities (97 instances, solely 8.76% of the set).

3.3 Findings on the enclosed paragraph types

As shown above (see 2.1), the narrow P-theme group enclose the two major paragraph types, viz. those featuring a stable P-theme and those where the P-theme is unfolded. However, their representation in the corpus turned out to be strikingly unbalanced, the former clearly outnumbering the latter. More specifically, the stable P-theme paragraphs covered over three-fourths of the narrow P-theme paragraphs (847 altogether; 76.44%), while the unfolded P-theme counterparts less than a fourth (261 in all; 23.56%). It follows that among the narrow P-theme group, the former pattern appears to epitomize the prototype.

Furthermore, each of the two types of paragraphs involved in the narrow P-theme set may be examined from the viewpoint of register preferences. Generally speaking, the researcher notices their rather uneven distribution across the corpora as well as subcorpora. Within the group of stable P-theme paragraphs, the most frequent turned out to be those detected in fiction instances (440 altogether, i.e. 51.95%). (It should be noted that more than a half of this number was composed of paragraphs featuring direct speech). Next came specimens drawn from journalism (315 in all, i.e. 37.19%), and those extracted from the academic corpus (92 altogether, i.e. 10.86% of the set).

By comparison, within the group of unfolded P-theme paragraphs the situation was found to be almost reversed. In the fiction corpus the share of the paragraphs turned out to be the lowest of all (36 instances, i.e. 13.79% of all), in journalism it increased (76 cases, i.e. 29.12% of the set), but it was striking in the academic corpus (149 paragraphs, i.e. 57.09% of all).

For a more accurate picture, attention will be turned also to the distribution of the two types of paragraph across the individual subcorpora. The findings are provided in Table 1.

	S abs.	S %	U abs.	U %	N abs.	N %
AS	39	4.60	105	40.23	144	12.99
AH	53	6.26	44	16.86	97	8.76
JN	244	28.81	43	16.48	287	25.90
JM	71	8.38	33	12.64	104	9.39
FCH	200	23.61	20	7.66	220	19.86
FA	240	28.34	16	6.13	256	23.10
TOT	847	100.00	261	100.00	1,108	100.00

Table 1: Distribution of narrow P-theme paragraphs across subcorpora

Key:	JM: journalistic subcorpus of magazine texts
Abs.: absolute value	JN: journalistic subcorpus of newspaper texts
AH: academic subcorpus of humanities	N: narrow P-theme paragraphs
AS: academic subcorpus of sciences	S: stable P-theme paragraphs
FA: adult fiction subcorpus	TOT: total
FCH: children's fiction subcorpus	U: Unfolded P-theme paragraphs

As Table 1 shows, the representation of the stable P-theme paragraphs in the individual subcorpora also proved to be inequitable. In fact, the paragraphs in question flourished in three subcorpora, but were rather rare in the remaining ones. More specifically, they thrive in the newspaper subcorpus (244 paragraphs, i.e. 28.81% in all) and in the adult fiction subcorpus (240 cases, i.e. 28.34% in all). Interestingly enough, these two subcorpora alone formed nearly two thirds of the instances in question. Children's fiction also proved to be a relatively prolific source of the paragraphs under examination (200 specimens, i.e. 23.61%). However, in all the other subcorpora, the share of the paragraphs under scrutiny was found to be comparably negligible. More specifically, the magazine subcorpus provided 71 instances (8.38%), subcorpus of humanities comprised 53 examples (6.26%) and the academic subcorpus of natural sciences included only 39 cases (4.60%).

The unfolded P-theme paragraphs were the most frequent of all in the subcorpus of academic sciences (105 cases, 40.23%). Somewhat lower share was exhibited by the subcorpus of academic humanities (44 instances, 16.86%) and by the newspaper subcorpus (43 specimens, 16.48%). Much rarer turned out their incidence in the magazine subcorpus (33 cases, 12.64% in all). Both the fiction subcorpora displayed the lowest values of all, with the children's fiction (20 instances, 7.66% in all) somewhat prevailing over its adult fiction counterpart (16 cases, 6.13%).

From the foregoing discussion it follows that the overall representation of

narrow P-theme paragraphs is largely determined by the decisive share of the stable P-theme gravitational field. The above data also suggest that at least in the corpus investigated, a clear affinity emerges between the fiction samples and the narrow P-theme paragraphs. Conversely, the relatively low incidence of these instances in the academic set appears to indicate that the pattern is much too constrained for the register's needs and therefore it yields to the broad P-theme paragraphs instead. Presumably, the academic corpus favours somewhat greater instability in paragraph build-up, the fiction corpus prefers higher stability, and the journalistic corpus seems to strike the right balance regarding the distribution of narrow and broad P-theme paragraphs.

3.4 Central vs. non-central narrow P-theme paragraphs

The above findings may be further complemented by a quick look at the proportions of the central and the non-central paragraphs in both the gravitational fields. In the present framework, non-central paragraphs in principle satisfy the respective defining criteria posited for narrow P-theme paragraphs, although they display their marginal weakening on formal and/or content grounds.

In contradistinction to the central specimens adduced in 2.2 above (I-X), the following examples epitomize non-central instances. Whereas (XI) is drawn from the stable P-theme field, (XII) is extracted from the unfolded P-theme field.

- (XI) 1. *Impressive as all this is, the grand claims made by some at the euro's birth look overblown.* 2. *It has not turned Europe into a dynamic economic powerhouse.* 3. *Nor can it, if governments continue to shy away from microeconomic reforms.* 4. *And, while the euro has taken the Deutschemark's place as the second commonest currency in official foreign-exchange reserves, it still has only one-fifth of the dollar's share.* (J6, 7)
- (XII) 1. *Other attractive fuel options include electrical power (even though electricity generation itself usually yields carbon dioxide) and hydrogen.* 2. *Hydrogen can be produced from water, again using electrical power.* 3. *When burned, the hydrogen simply re-converts into water.* 4. *Both BMW and Mercedes-Benz are developing hydrogen-powered cars. (The Observer 1989, p.47).* 5. *Ironically, hydrogen fuel may be most easily and benignly available in tropical countries where electricity can be generated by solar power.* (A6, 75)

In the research done, of the 1,108 narrow P-theme instances collected, 686 specimens came from the central zones, and 422 were non-central. (Naturally, the latter figure combines the transitional and the peripheral zones of the two gravitational fields.) This means that nearly two thirds of the paragraphs in question embodied the build-up patterns in their canonical forms. More specifically, 61.91 per cent of the central narrow P-theme paragraphs were contrasted with 38.09 per cent of their non-central counterparts.

Naturally, centrality and marginality can be studied also with respect to register distribution. The findings are overviewed in Table 2.

	A abs	A %	J abs	J %	F abs	F %	TOT abs	TOT %
C	63	26.14	302	77.24	321	67.44	686	61.91
NC	178	73.86	89	22.76	155	32.56	422	38.09
TOT	241	100.00	391	100.00	476	100.00	1108	100.00

Table 2: Distribution of central and non-central narrow P-theme paragraphs across corpora

Key:

A: academic corpus

Abs: absolute value

C: central narrow P-theme paragraphs

F: fiction corpus

J: journalistic corpus

NC: non-central narrow P-theme paragraphs

TOT: total

As Table 2 shows, in the fiction corpus, there were 321 central paragraphs, displaying a narrow P-theme, as against only 155 non-central ones. Clearly, in the fiction corpus, paragraphs were discovered to favour the two distinct types of narrow P-theme paragraphs in their prototypical formats, as 67.44 per cent of fiction paragraphs came from central zones.

Conversely, academic sources gave preference to the two paragraph types in their non-canonical manifestations. Indeed, against sixty-three central paragraphs with a narrow P-theme, there were altogether 178 non-central ones. This suggests that when it comes to the build-up of narrow P-theme paragraphs, the academic corpus, unlike its fiction counterpart, connotes the relatively greatest instability of all. More specifically, altogether 73.86 per cent of the academic paragraphs in question were found to accommodate some modifications and enrichments of constructional patterns. Presumably, for the relatively long paragraphs in academic texts, the patterns are too tight and absolute adherence to them would be felt as limiting.

Last of all, in journalism, among the narrow P-theme paragraphs, as many as 302 central specimens (i.e. 77.24%) may be correlated with their 89 non-central counterparts. These data testify unequivocally to the fact that in journalism the

majority of paragraphs under scrutiny epitomized paragraph types adhering strictly to the respective build-up patterns. Nevertheless, this tendency detected might have been influenced crucially by the relative brevity of the typical newspaper paragraphs displaying a narrow P-theme, offering arguably little room for various departures, inconsistencies, enrichments, etc.

3.5 Narrow P-theme paragraphs featuring direct speech

A sizeable proportion of narrow P-theme paragraphs was constituted by specimens featuring portions of direct speech (DSp). Indeed, of the 527 DSp-paragraphs collected, there were in all 444 specimens marked by a narrow P-theme. This amount corresponds to 84.25 per cent of all the DSp-specimens assembled.

Within the stable P-theme field, there were 409 DSp-specimens, which accounted for a striking 48.29 per cent of all. Among the unfolded P-theme set, the incidence of DSp-paragraphs was comparably much lower. More specifically, 35 instances were extracted, corresponding only to 13.40 per cent of the field. The following examples feature portions of direct speech. Whereas (XIII) illustrates a stable P-theme paragraph in question, (XIV) exemplifies its unfolded P-theme counterpart.⁹

- (XIII) 1. *Mary* said: 'I didn't clap when that Sally Mahonia won. 2. I didn't clap at all.' 3. And then, tired out from scratching her chest and eating cake and wanting Pearl to be recognized as the Most Beautiful Baby in Swaithey, *she* fell asleep in Irene's lap. (F6, 57)
- (XIV) 1. *The families directly affected* admit they are anxious to find a cause. 2. At the same time *some* are wrestling with anger and self-blame on top of their loss. 3. *Gill Callar* wonders whether the family should have moved from Berkshire to Heston seven years ago. 4. "We used to come down here for family holidays and we all loved it. 5. Emma was very happy at school but after a while she started complaining of pains in her legs. (J1, 62)

It should be remarked, however, that the paragraphs marked by direct speech were not spread across the register corpora evenly. Direct speech appeared most readily in fiction, less so in journalism, and least frequently in academic writing. In particular, of the paragraphs displaying direct speech, 304 specimens came from fiction (44.06% of the fiction corpus), 153 from journalistic sources (22.17% of the journalistic corpus) and 70 were extracted from academic texts (10.14% of the academic corpus).

Since in the academic corpus the instances in question served mostly

intertextuality purposes and were in many respects rather peripheral, they may be disregarded for the sake of greater comparability of data. Focussing on the journalism and fiction corpora alone, there were in all 457 paragraphs involving direct speech. This group was composed of 392 narrow P-theme instances, and sixty-five broad P-theme cases. It follows that among the direct speech paragraphs in the more homogeneous set, the narrow P-theme paragraphs account for 85.78 per cent of all, relegating 14.22 per cent of specimens to the broad P-theme pattern. These data confirm a conspicuous affinity between narrow Theme paragraphs and specimens featuring direct speech rather convincingly.

3.6 Narrow P-theme paragraphs in terms of length

The narrow P-theme paragraphs comprised altogether 1,108 instances. These in turn enclosed 3,298 utterances/4,127 (main) clauses in all. It follows that the mean paragraph length was 2.98 utterances/3.72 clauses. Such values, however, fell far below the overall corpus average (3.6 utterances/4.49 clauses). Therefore at least in the corpus investigated, the narrow P-theme set connotes rather short paragraphs.

Much the same tendency holds true even within individual gravitational fields. Regarding the stable P-theme gravitational field, the central paragraphs were on average 2.16 utterances/2.67 clauses in length, those in the transitional zone turned out to be somewhat longer (2.79 utterances/3.53 clauses on average), rendering the peripheral ones unequivocally the longest of all in the field (with values 4.74 utterances/6.33 clauses). Thus of the three areas, only the specimens in the peripheral zone exceeded the overall research corpus average (i.e. 3.6 utterances/4.49 clauses) in terms of paragraph length.

Similarly, in the unfolded P-theme gravitational field, the central zone also embraced the shortest paragraphs (2.95 utterances/3.62 clauses), the transitional ones proved to be somewhat longer (3.81 utterances/4.55 clauses), and the peripheral zone displayed again the longest specimens of all (5.05 utterances/6.11 clauses). It should be remarked that whereas the central unfolded P-theme paragraphs fell, in length terms, below the overall corpus average, both the non-central zones included instances exceeding the overall corpus mean lengths.

Moreover, when the individual zones all over the narrow P-theme paragraphs are arranged according to the growing paragraph mean lengths, the shortest specimens of all are detected among the central (2.16 utterances/2.67 clauses) and transitional (2.79 utterances/ 3.53 clauses) stable P-theme paragraphs. The central unfolded P-theme instances come next (2.95 utterances/3.62 clauses).

Much longer cases originate in the peripheral zone of the stable P-theme field (3.6 utterances/4.49 clauses). Still longer turn out to be transitional specimens marked by the unfolded P-theme (3.81 utterances/4.55 clauses). Finally, the longest instances of all come from the peripheral zone of the paragraphs where the P-theme is unfolded (5.05 utterances/6.11 clauses).

It is presumably by no means coincidental that the shortest paragraphs were found to display the top degree of stability. Conversely, the longest specimens among those marked by a narrow P-theme exhibited the extremest measure of constructional instability of all.

Last but not least, paragraph length is affected by the register. Indeed, irrespective of the paragraph type, significant differences in length were detected across the individual research corpora. Generally speaking, the journalistic corpus embraced the shortest paragraphs of all (2.91 utterances/3.37 clauses). In fiction, the instances proved to be somewhat longer, and with their mean length of 3.21 utterances/4.49 clauses, they came the closest to the overall corpus average. In the academic corpus, the specimens turned out to be convincingly the longest of all, embracing on average 4.67 utterances/5.70 clauses.

3.7 Register-specific tendencies pursued in the narrow P-theme paragraphs

Regarding the remarkable paragraphing tendencies detected in individual register corpora and subcorpora, in fiction, the specimens generally favoured the top measure of constructional stability. Not only was the frequency of narrow P-theme paragraphs the most striking of all, but among them, the stable P-theme paragraphs were dominant. Moreover, the narrow P-theme paragraphs featured were usually central, adhering strongly to the constructional patterns.

Furthermore, of all the scrutinized corpora, fiction was marked by the top degree of paragraph length heterogeneity. One cause of paragraph length variability was the intertwining of paragraphs with a share of direct speech and those devoid of direct speech portions. The research done revealed a striking length discrepancy between paragraphs in these groups, the former being marked by conspicuous brevity. It should be stressed, however, that the dichotomy represents solely one factor. Indeed, fiction is much too creative, and variegated to be considered a typical 'register'.

This stylistic variety was generally marked by a limited number of the thematic DSs on the scene of discourse, each, as a rule, foregrounded at a time. Narrative passages clearly prevailed over descriptive ones. As observed already by Červenka (1982), narrative texts tend to display paradigmatic arrangements rather willingly.

Direct speech also ranked among the prominent features, enlivening the

prose. Since the present research discovered a clear affinity between the narrow Theme paragraphs and paragraphs displaying direct speech, it is by no means surprising that fiction proved to be an ample source of narrow Theme paragraphs, particularly stable Theme ones.

Similar to fiction, journalism texts were also prone to display a restricted number of thematic DSs. This feature, however, appears pragmatically related to the conspicuous space constraints imposed on these texts. Characteristically, the journalism samples under investigation favoured central specimens and exhibited a tendency towards paragraph brevity, which proved to be even more pronounced in the newspaper subcorpus. Furthermore, they were marked by the greatest homogeneity in paragraph length of all. Naturally, the prototypical column arrangement turned even very short passages into eye-paragraphs, producing nearly a paragraphing rhythm. Consequently, many a notional paragraph was realized in two orthographic ones. Not surprisingly, the register ranks among the relatively closed ones. On the one hand, the brevity of the paragraphs fails to provide room enough for unfolding the Theme properly; on the other hand, the space limits imposed on the texts as such call for acceleration in internal thematic turnover. These mutually conflicting tendencies result in a relatively even share of the two constructional patterns in the narrow Theme set. What is more, of all the corpora under investigation, this register even strived to strike the finest balance between the incidence of the narrow and the broad P-theme paragraphs.

In journalistic samples, meticulous scrutiny is mostly missing. Stories should be succinct, catchy and swift to read. Therefore, the graphic aspect appears to gain in significance. The surveyability of these paragraphs is enhanced by such devices as headlines, subheadlines, leads, columns, etc. Frequently, the P-theme exponent is advertised at the very onset of a paragraph. Furthermore, linear arrangement of paragraphs is reclassified into a means of hierarchizing them (for the broader interpretation of iconicity, see e.g. Enkvist 1991).

As far as academic texts are concerned, they featured narrow P-theme paragraphs rather sparingly. Generally speaking, the register displayed the longest paragraphs of all. The characteristically long paragraphs betray the author's ambition to achieve an accurate, objective, and balanced account of the analysed phenomena. Their considerable space could comfortably accommodate broad P-theme patterns, where the focus is put on contextualizations, interrelations of various DSs, and at times, also on dynamic treatments and interpretations. Indeed, broad P-theme paragraphs represent a happy choice for argumentative and expository purposes and are suitable where multifaceted and multiaspectual treatments of facts are desirable. These tendencies are aptly matched with various graphic means, such as numbering, lettering, subtitles, etc. Such means,

enhancing clarity, are useful and justified, for lengthier paragraphs are generally considered relatively more demanding for processing.

Of the narrow P-theme specimens, the register preferred the paragraphs marked by a relatively higher measure of constructional instability, i.e. those where the P-theme is unfolded. Naturally, this pattern suits subtle analysis and classification. Clearly, unfolding Themes serve better the meticulous scrutiny and structured accounts than the limiting stable Themes, as the latter would render the content rather flat and devoid of subtle hierarchy.

Significantly, academic texts were prone to employ the narrow P-theme paragraphs from both the transitional and the peripheral zones of the two relevant gravitational fields, as their central counterparts would hardly accommodate all the needed evaluative comments, departures, contextualizing clues, arguments, asides, notes, crossreferences, etc.

Thus, the academic subcorpus clearly favoured patterns marked by higher degrees of build-up instability, a more pronounced thematic turnover, and constructional diversity, since such patterns allowed for comprehensive, accurate and meticulously structured treatments of phenomena.

It seems that the paragraphing tendencies detected in the individual register corpora and subcorpora reflect primarily the differences in the subject matter dealt with. Among others, the particular subject matter has a bearing on the significance assigned to the various language functions, on the types of units employed in thematic functions or on the diverse conventions adhered to.

4 Conclusion

This paper explores the narrow P-theme paragraphs in authentic English texts. These paragraphs, marked by the top degrees of stability in their thematic build-up, are prone to induce coherence rather strongly. They foreground a single major Thematic discourse subject. Its explicit exponent, which is typically launched through the initial Topic Sentence, motivates the U-thematic functions in all the paragraph utterances. The thematic progressions are chiefly paradigmatic and the cohesive links interlacing the thematic elements establish, but never exceed, the narrow paradigm (involving identity, equivalence and inclusion). The narrow P-theme supertype encloses two major paragraph types, viz. those featuring a stable P-theme and those where the P-theme is unfolded.

In a corpus of contemporary British English texts, these paragraphs were marked by their prototypical brevity, displayed a remarkable tendency to adhere to the canonical, rather than marginal, patterns and showed their willingness to feature portions of direct speech. In the corpus, the stable P-theme specimens unequivocally epitomized the prototype, since they by far outnumbered their

unfolded P-theme counterparts. The thematic progressions typically fell in the continuous theme type. Derived themes were generally rather marginal and were confined to the unfolded P-theme subgroup. As for the register preferences, the narrow P-theme paragraphs flourished particularly in fiction texts, but were rather reluctant to appear in academic samples. Among the subcorpora, their highest frequency was attested in newspaper texts and in adult as well as children's fiction. The various paragraphing tendencies were shown to be related to the subject-matter differences existing between the registers explored.

Symbols and abbreviations

- A1-6: sources composing the academic corpus
- DS: Discourse Subject
- DSp: direct speech
- F1-6: sources composing the fiction corpus
- FSP: Functional Sentence Perspective
- J1-6: sources composing the journalistic corpus
- P-theme: Paragraph Theme
- U-theme: utterance theme

Notes

¹ To distinguish between the two homonymous terms of "THEME", in what follows (outside citations), we shall reserve the capital-preceded "themes" – i.e. (Textual) Theme and its conspicuous variety, P-theme (Paragraph Theme) as interpreted on the textual, hierarchically superior level. The non-capitalized "theme", on the other hand, will label its counterpart delimited on the hierarchically inferior FSP level and contrasted with the rheme (see, e.g., Firbas 1992; Adam 2006). It should be noted that the non-capitalized "theme" roughly corresponds to "Topic" in Hajičová and Sgall (2004).

² In the present approach, the content aspect of the Theme, constituting the so-called Thematic area, is conceived of as enclosing in all three distinct layers arranged from the broadest to the narrowest, to resemble a kind of a pyramid. The lowest and broadest layer, which is simultaneously the most diffuse of all, follows from the overall communicative framework. It corresponds to all the given elements of the speech event. The central layer embraces a number of hierarchized, closely interrelated and regularly co-occurring elements organized as a cognitive structure, or a content frame. The third, the most restricted one of all, though also potentially available (at least) in (some) texts, embodies some of its most conspicuous or foregrounded elements, or else, its dominant Discourse Subjects.

³ "As discourse subject (DS) I treat anything – be it an object, a group or class of them, a quality, state, process, action, circumstance, event, episode, and the like – that the speaker has just in mind when applying a nominating (or deictic) unit in the process of text production in order to introduce/present/mention/re-introduce/recall something." (Daneš 1989: 24)

⁴ The term "thematic progressions" is adopted in the sense defined by Daneš (1974: 114), i.e. the choice and ordering of utterance themes, their mutual concatenation and hierarchy, as well as their

relationship to the hyperthemes of the superior text units (such as paragraph, chapter,...) to the whole text, and to the situation.” Their comprehensive classification is provided in Daneš (1989: 25-26).

⁵ In this paper, cohesive links are conceived of as arranged on a tentative scale running from full identity, via equivalence, inclusion, all the way to parallelism and contiguity. Identity and equivalence do not involve any change in reference from their antecedents, although the latter is characterized by change of other modes of meaning; inclusion represents referential overlap; parallelism presupposes referential commensurability with respect to a common reconstructable superordinate unit; and contiguity, the least clear-cut type of all, involves some relation of mutual expectability, co-occurrence, semantic relatedness, association, etc.

Somewhat modifying the concept of thematic paradigm introduced by Červenka (1982), we distinguish between the narrow thematic paradigm on the one hand, and the broader thematic paradigm on the other. The former is taken to include the proportion of identity, equivalence and inclusion among the thematic elements related to a single Discourse Subject, while the latter is conceived of as incorporating all the remaining cohesive relations established between thematic units linked to a single Discourse Subject.

⁶ “Foregrounding” will be understood essentially in line with the Prague linguistic tradition, particularly with Mukařovský 1932 (2000). Mukařovský argues that the purpose of foregrounding is “to attract the reader’s (listener’s) attention more closely to the subject matter expressed by the foregrounded means of expression.” (ibid: 227) Foregrounding is intentional. It implies choice and is related to hierarchy. “The component highest in the hierarchy becomes the dominant. All other components, foregrounded or not, as well as their interrelationships, are evaluated from the standpoint of the dominant. The dominant is that component of the work which sets in motion, and gives direction to, the relationships of all other components” (ibid.).

⁷ The method of underlining adopted enhances the surveyability and transparency of our analysis, though at the expense of some inevitable simplification. It should be stressed that the perception of coherence comes out of an interplay of numerous other links and elements (including other thematic units both in the same and other distributional fields (see e.g. Svoboda 1968).

⁸ It should be noted that for the purposes of this paper no distinction is made between these terms.

⁹ It is well known that the rather indeterminate syntactic status of the reporting clause has prompted diverse linguistic treatments. In this paper, in line with Dušková et al. (1988), the DSp stretches in texts will be considered as special instances of content clauses. Thus, in the present research a reporting clause attached to the reported direct speech utterance counts as a single utterance. Since more reporting clauses may theoretically appear in a single paragraph, the remaining direct speech utterances are considered to have the reporting clauses ellipted.

References:

- Adam, M. (2006) *Functional Macrofield Perspective. (A Religious Discourse Analysis Based on FSP)*. Brno: MU.
- Červenka, M. (1982) ‘Narration and description from the standpoint of functional sentence perspective.’ In: P. Steiner, P. Červenka, M., Vroon, R. (eds) *The Structure of the Literary Process*. Amsterdam/Philadelphia: John Benjamins Publishing Co., 15-44.
- Daneš, F., (ed.) (1974) *Papers on Functional Sentence Perspective*. Praha: Academia.
- Daneš, F. (1989) ‘Functional sentence perspective and text connectedness.’ In: Conte, M.-E., Petőfi, J. S., Sözer, E. (eds) *Text and Discourse Connectedness* Amsterdam/Philadelphia: John Benjamins Publishing Co., 23-31.

- Daneš, F. (1994) 'Odstavec jako centrální jednotka tematicko-kompoziční výstavby textu (na materiále textů výkladových).' *Slovo a slovesnost* 55, 1-17.
- Daneš, F. (1995) 'Paragraph – a central unit of the thematic and compositional build-up of texts.' In: Wårwik, B., Tanskanen, S. K., Hiltunen, R. (eds) *Organization in Discourse. Proceedings from the Turku Conference*. Turku: University of Turku. 29-40.
- Dušková, L. et al. (1988) *Mluvnice současné angličtiny (na pozadí češtiny)*. Praha: Academia.
- Enkvist, N. E. (1991) 'Discourse strategies and discourse types.' In: Ventola, E. (ed.) *Functional and Systemic Linguistics. Approaches and Uses*. Berlin/New York: Mouton de Gruyter. 5-22.
- Firbas, J. (1992) *Functional Sentence Perspective in Written and Spoken Communication*. Cambridge: Cambridge University Press.
- Geluykens, R. (1991) 'Information flow in English conversation: A new approach to the given – new distinction. In: Ventola, E. (ed.) *Functional and Systemic Linguistics. Approaches and Uses*. Trends in Linguistics, Studies and Monographs 55, Berlin/New York: Mouton de Gruyter. 141-167.
- Hajičová, E., Sgall, P. (2004) 'Degrees of contrast and the Topic-Focus Articulation.' In: Steine, A. (ed.) *Informational Structure. Theoretical and Empirical Aspects*. Berlin/New York: Walter de Gruyter. 1-13.
- Mathesius, V. (1942/1982) 'Umění psát odstavce. Řeč a sloh.' In: *Jazyk, kultura a slovesnost*. Praha: Odeon, 92-146.
- Mukařovský, J. (2000) 'Standard language and poetic language (1932).' In: Burke, L. Crowley, T., Girvin, A. (eds) *The Routledge Language and Cultural Theory Reader*. Chapter 23, London/New York: Routledge, 2000. 225-230.
- Pípalová, R. (2008) 'On narrow and broad P-theme paragraphs.' *Linguistica Pragensia* XVIII/1, 37-51.
- Pípalová, R. (forthc.) *Thematic Organization of Paragraphs and Higher Text Units*. Praha: UK PEDF.
- Svoboda, A. (1968) 'The hierarchy of communicative units and fields as illustrated by English attributive constructions.' *Brno Studies in English* 7, 49-85.

Sources

- A1** Norwich, J. J. (ed.) (1991) *Britain's Heritage*. London: Kingsfisher Books.
- A2** Watkin, D. (1990) *English Architecture. A Concise History*. Reprinted version. London: Thames and Hudson.
- A6** Yearley, S. (1991) *The Green Case. A Sociology of Environmental Issues, Arguments and Politics*. London/New York: Routledge.
- J1** www.guardian.co.uk accessed on 25/11/03; 28/11/03.
- J2** www.independent.co.uk accessed on 29/11/03.
- J4** www.newscientist.com accessed on 6/1/02.
- J6** www.economist.com accessed on 3/1/02.
- F1** Carruth, J. (1991) *Treasury of Fairy Tales. Goldilocks and the Three Bears*. London: Award Publications.
- F2** Rowling, J. K. (1999) *Harry Potter and the Prisoner of Azkaban*. London: Bloomsbury.
- F6** Tremain, R. (1992) *Sacred Country*. A Sceptre Paperback. London: Hodder and Stoughton.