CLINICAL LINGUISTICS & PHONETICS, 2003, VOL. 17, NO. 4–5, 355–363

Prosodic and lexical-syntactic aspects of the therapeutic register

LI SHENG, KARLA K. MCGREGOR and YI XU

Department of Communication Sciences and Disorders, Northwestern University, IL, USA

(*Received 10 September 2002; accepted 9 November 2002*)

Abstract

The current study examined features of speech addressed to younger versus older and impaired versus normal listeners and explored factors that might motivate speech modification in the therapeutic register. Ten speech-language clinicians in training were engaged in a simulated storytelling task and produced narrations to four imaginary listener groups. Results revealed that clinicians modified their speech at multiple levels according to both the age and the diagnostic status of their listeners. As a group, the clinicians demonstrated decreased speaking rate, elongation of pauses, and use of shorter and less complex sentences when speaking to young children and individuals with language impairments. Despite these group trends there was notable individual variability between clinicians. Findings from this constrained experimental context may be extended to real life clinician-client interactions and have implications for therapeutic success.

Keywords: Therapeutic register, age, diagnostic status, intervention.

Introduction

Motherese, a simplified register of speech directed to children, has been extensively studied in the past three decades. Various studies have documented the many distinctive features of motherese, examined the variability of this register across cultures, age groups, and social classes, and discussed the role of this specialized input in language acquisition. A range of variables have been investigated in these studies, which addressed the prosodic, lexical, grammatical, and pragmatic aspects of motherese. Motherese has been characterized as a clear, simple, and repetitive way of speech (Snow, 1994). More specifically, motherese is short in utterance length, slow in tempo, high in pitch, exaggerated in intonation contour, and simplified in lexical choice and syntactic structure (for more information, see Snow and Ferguson, 1977; Gallaway and Richards, 1994). The findings that motherese is

Address correspondence to: Li Sheng, 2240 Campus Dr., Northwestern University, Evanston, IL 60208, USA. e-mail: lil-sh@northwestern.edu

Clinical Linguistics & Phonetics
ISSN 0269-9206 print/ISSN 1464-5076 online © 2003 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/0269920031000079985

common in many cultures and is demonstrated by parents, adult strangers, and even older siblings have motivated the argument in some researchers that it is a species-specific behaviour preprogrammed in humans (e.g., Papoušek, Papoušek and Haekel, 1987).

Despite the large body of research on child-directed speech, studies with a focus on atypical learners lag far behind (Conti-Ramsden, 1994). Further, occupation-related differences in child-directed speech are also understudied (Shute, 1987). A handful of studies have hinted about the use of a distinctive speech register in the therapeutic setting (Panagos, Bobkoff and Scott, 1986; Reuvers and Hargrove, 1994), but no firm conclusions have been reached. In the present study, we further examined the features of the *therapeutic register*, namely, the speech of speech-language pathologists during clinical settings. Because of the possibility of subtypes of therapeutic register for subtypes of impaired population, we limited our investigation to speech directed to children with language impairments. We hypothesized that age as well as the diagnostic status of the listener would motivate the clinicians towards speech modification at multiple levels, such as prosody, lexical diversity and syntactic complexity. The diagnostic status factor would play a role over and above the age factor so that we should see two independent main effects in our analyses. We put the hypothesis under rigorous test by engaging the clinicians in a book-reading task with imagined child listeners instead of real child interactants. This design was based on the premise that modifications found in a more demanding condition could be generalized to real life clinician-client interactions.

Method

Subjects

Ten female, native speakers of American English participated in this study. All subjects were master or doctoral students of Speech-Language Pathology at Northwestern University, USA. Subjects' experience with clinical populations and with children was not controlled, but relevant inquiries were made during the debriefing. At the time of participation, six subjects had approximately half a year of clinical training, three subjects had roughly 1 and a half years of clinical training and one subject had more than 5 years of clinical practice.

Material

We used a short story titled *Just A Pig At Heart* from a wordless picture book by Mayer (1974) to elicit narratives from the subjects. The story was about a female and male pig getting dressed up to go out and consisted of 14 pictures of simple black and white drawings. The pictures were reprinted onto letter size papers and displayed on a poster board. On some pages, a noun referring to a pictured object was labelled. A total of 10 nouns were selected and printed on the story. Subjects were asked to include the 10 printed words in their narrations.

Four exemplar language profiles representing four listener groups were concocted to create the storytelling conditions. Each profile consisted of a figure depicting an individual's percentile scores on several standardized language tests with reference to norms and a short descriptive text to assist interpretation of the individual's performance (see the Appendix for a sample of the mock profiles). The

four listener groups differed in terms of age and diagnostic status so that two groups were 3-year-olds and the other two were 12-year-olds. Among each age group, half were said to be normally developing and half were language impaired. The profiles created for the older and younger impaired listeners demonstrated equivalent severity of impairment and the same was true for the normal groups.

Procedure

The experiment took place in a comfortably furnished laboratory playroom and lasted for an average of 40 minutes. To ensure the validity of the data, the real purpose of the study was first masked to the subjects. Subjects were told that the researchers were interested in children's story comprehension and they were invited to record stories for the children. After the masking procedure, subjects were left alone in the playroom to review the story and practice telling the story. Following the practice, the researcher presented the language profile of a typical member from the first listener group and the subject was given time to review the profile and ask questions. When the subject was ready, the researcher withdrew from the playroom and started recording. Upon completion of the first recording, the researcher went back to the playroom and presented the next profile to the subject and the above procedures were repeated. Each subject underwent four different conditions. The order of presentation of conditions was counterbalanced across subjects. Half the subjects narrated stories to the two 3-year-old groups first, half to the 12-year-old groups first. In addition, half the subjects received a normal-impaired order and half an impaired-normal order for the same-age groups. Following data collection, the subject was debriefed. The subject was asked to comment on how she perceived the task and whether she made any adjustments in her speech. Only one subject indicated suspicion of being tested whereas all the remaining nine subjects said they made modifications inadvertently as a result of clinical training.

Acoustic analysis

A total of 40 narrations were audio-recorded using a Marantz PMD-500 professional tape recorder and a Samson lavaliere microphone. Each narration was approximately one to three minutes in length. For the purpose of acoustic analysis, each narration was digitized individually using SoundEdit 16, with a sampling rate of 22 KHz. The first minute of continuous speech within each narrative sample was included in the acoustic analysis. Acoustic data were analysed using Praat 4.0.11 (Boersma, 2001) and Matlab 5.3. Dependent measures included utterance duration, pause duration and speaking rate. An utterance was acoustically defined as continuous speech bounded by silent intervals of 300 milliseconds (ms) or longer (Fernald and Simon, 1984). A pause was longer than 300 ms and shorter than 3000 ms (Stern, Spieker and Barnett, 1983). Speaking rate was calculated by dividing the total number of syllables by the total speaking time, excluding pauses (Fernald and Simon, 1984).

Narrative sample analysis

Narrative samples were transcribed in full length according to SALT conventions. Utterances were segmented following both prosodic and grammatical cues. Several

subjects used 'and' very frequently to connect utterances. Continuously conjoined utterances like these were arbitrarily segmented after one conjunction. Word level codes were inserted into the narrative samples to denote any use of nouns (including common and proper nouns). Utterance level codes were made to mark the use of passive structures, compound sentences (sentences with coordination) and complex sentences (sentences with one or more subordinated clauses). The following dependent measures were obtained directly or derived from the transcripts and codes: number of different words (NDW), number of total words (NTW), type token ratio (TTR), noun types, noun tokens, TTR for nouns, number of total utterances, mean length of utterance in morphemes (MLU-M), and numbers of simple sentences and passive structures. In addition, we divided the lexical measures by the total number of utterances to examine word uses in natural units of speech.

A second person independently transcribed and coded eight randomly selected samples (20% of the data). Point-to point agreement on utterance boundaries averaged 92.5% and ranged from 86.5% to100%. Agreement on total morphemes averaged 98.5% and ranged from 97.4% to 99.5%. Agreement on the type of morphemes ranged from 98.7% to 100% and averaged 99.4%. Point-to-point inter-coder reliability on word level codes averaged 90.7% and ranged from 86.4% to 96%. Agreement on utterance level codes ranged from 70% to 90.1% and averaged 82.7%. All disagreements were resolved by consensus.

Results

Repeated measures ANOVA was run on all dependent measures with listener age (3-year-old, 12-year-old) and diagnostic status (impaired, normal) as the within subject variables. Results on the acoustic measures were presented in table 1. When utterance duration was the dependent measure, the differences between the conditions failed to reach significance, although there was the numerical tendency for utterances directed to the 12-year-olds and the normal listeners to be longer than those directed to the 3-year-olds and the impaired children. ANOVA on pause duration revealed a main effect of diagnostic status (F (1,9) = 5.08, $p = 0.05$, $\eta^2 = 0.36$). As predicted, clinicians inserted longer pauses when the hypothetical listeners were individuals with language impairments. There was no main effect for age and no age x diagnostic status interaction. When the dependent measure was syllables per second, the results were a main effect of age (F (1,9) = 7.03, $p = 0.03$, $\eta^2 = 0.44$), and a marginal main effect of diagnostic status (F (1,9) = 3.16, $p = 0.05$

Table 1. *Mean values (SD) for the acoustic measures (in seconds or syllables per second)*

	Impaired 3-year-olds	Normal 3-year-olds	Impaired 12-year-olds	Normal 12-year-olds
Utterance duration	2.224 (0.545)	2.344 (0.461)	2.293 (0.459)	2.447 (0.573)
Pause duration[2,*]	0.941 (0.187)	0.864 (0.231)	0.957 (0.193)	0.903 (0.124)
Speaking rate[1,*;2,#]	3.879 (0.514)	4.089 (0.502)	4.144 (0.576)	4.383 (0.352)

Notes: 1 = a significant effect for age, 2 = a significant effect for diagnostic status. *Significant at the 0.05 level for a two-tailed test. #Significant at the 0.05 level for a one-tailed test.

(one-tailed), $\eta^2 = 0.26$). Clinicians spoke faster to the 12-year-olds and the normal children. No interaction was found for speaking rate.

Table 2 listed the mean values for measures of lexical diversity and syntactic complexity. We first examined lexical diversity per story. Repeated measures ANOVA on NDW, the measure for overall lexical diversity, revealed no significant main effect or interaction. When NTW was the dependent measure, there still was not a main effect. Nor did the interaction effect reach significance (F (1,9) = 3.74, $p = 0.09$). The analysis on TTR resulted in a significant age effect (F (1,9) = 3.67, $p = 0.04$ (one-tailed), $\eta^2 = 0.29$) and no other effects. The TTR values were lower for the 3-year-old groups than for the 12-year-old groups. Analyses on the types of nouns in the entire sample revealed a significant effect of diagnostic status (F (1,9) = 3.37, $p = 0.05$ (one-tailed), $\eta^2 = 0.27$) and no other effect. Consistent with our hypothesis, a larger variety of nouns was used with the normal children. When noun tokens was the dependent variable, no main effect was found, neither was there an interaction (F (1,9) = 4.31, $p = 0.07$). The type and token ratio on nouns also did not differentiate the conditions from each other.

We next calculated lexical diversity per utterance. ANOVA on NDW per utterance resulted in an effect of age (F (1,9) = 9.56, $p = 0.01$, $\eta^2 = 0.52$) and an effect of diagnostic status (F (1,9) = 15.37, $p = 0.004$, $\eta^2 = 0.63$). These main effects were also qualified by a significant interaction effect (F (1,9) = 4.98, $p = 0.05$, $\eta^2 = 0.36$), with there being a sharper increase of different word usage across the two 12-year-old groups. Results for NTW per utterance (equalled to MLU in words) were identical to the above. There were an effect of age (F (1,9) = 10.36, $p = 0.01$, $\eta^2 = 0.54$), an effect of diagnostic status (F (1,9) = 13.98), $p = 0.005$, $\eta^2 = 0.61$), and an interaction effect (F (1,9) = 5.12), $p = 0.05$, $\eta^2 = 0.36$). The increase in total words per utterance

Table 2. *Mean values (SD) for the lexical and syntactic measures*

	Impaired 3-year-olds	Normal 3-year-olds	Impaired 12-year-olds	Normal 12-year-olds
NDW	96.8 (25.52)	104 (15.32)	101.5 (24.49)	112.9 (31.89)
NTW	241.7 (103.81)	251.6 (79.21)	227.2 (71.84)	252 (80.14)
TTR[1,#]	0.42 (0.06)	0.43 (0.06)	0.46 (0.04)	0.45 (0.02)
Noun type[2,#]	23.4 (5.27)	25.8 (5.65)	25.7 (5.81)	28.9 (9.52)
Noun token	43.3 (14.64)	44.9 (11.36)	43.8 (9.46)	49.5 (13.07)
Noun TTR	0.57 (0.09)	0.58 (0.07)	0.59 (0.09)	0.58 (0.06)
NDW per utterance[1,**;2,**;3,*]	4.62 (1.33)	4.86 (1.06)	5.20 (0.97)	6.04 (1.02)
NTW per utterance[1,**;2,**;3,*]	10.82 (2.43)	11.27 (1.77)	11.42 (1.87)	13.34 (2.15)
Noun type per utterance[1,**;2,*]	1.14 (0.37)	1.19 (0.21)	1.32 (0.25)	1.54 (0.32)
Noun token per utterance[1,**;2,**;3,*]	2.02 (0.64)	2.05 (0.42)	2.27 (0.52)	2.69 (0.60)
Total utterances	23.8 (13.29)	23.1 (9.07)	20.2 (6.23)	19.3 (6.65)
Simple sentences[1,#]	15.60 (8.87)	14.20 (6.75)	12.40 (4.12)	11.30 (5.44)
Passive structures	0.40 (0.52)	1.00 (0.94)	1.10 (1.20)	0.90 (0.57)
MLU-M[1,**;2,**;3,*]	12.00 (2.41)	12.57 (1.93)	12.71 (1.95)	14.83 (2.38)

Notes: 1 = a significant effect for age, 2 = a significant effect for diagnostic status, 3 = a significant interaction effect. *significant at the 0.05 level for a two-tailed test. **significant at the 0.01 level for a two-tailed test. #significant at the 0.05 level for a one-tailed test.

was smaller from the 3-year-old impaired to the 3-year-old normal groups than from the 12-year-old impaired to the 12-year-old normal groups. Analyses on noun types per utterance resulted in a main effect of age (F (1,9) = 10.39, p = 0.01, η^2 = 0.54), a main effect of diagnostic status (F (1,9) = 5.47, p = 0.04, η^2 = 0.38), and no interaction. More noun types were used with the older and the normal children. When noun tokens per utterance was the dependent measure, the results were a significant age effect (F (1,9) = 17.64, p = 0.002, η^2 = 0.66), a significant effect of diagnostic status (F (1,9) = 12.27, p = 0.007, η^2 = 0.58) and an interaction effect (F (1,9) = 5.58, p = 0.04, η^2 = 0.38). Again we saw a wider gap between the two 12-year-old groups than between the 3-year-old groups.

The last four variables in table 2 were measures of syntactic complexity. These numbers were based on the entire narrative samples. There was a tendency towards producing more utterances when narrating stories to younger children, although the age effect failed to reach significance (F (1,9) = 3.79, p = 0.08). Analyses on the number of syntactically simple sentences (with only one clause) resulted in a significant effect of age (F (1,9) = 3.73, p = 0.04, η^2 = 0.29) and no other effects. Not surprisingly, speech directed to younger children contained more simple sentences than that directed to older children. When the number of passive structures was entered into the analyses, no significant effect was found (the interaction effect also fell short of significance, F (1,9) = 3.69, p = 0.09). Finally, MLU in morphemes was analysed and there were a main effect of age (F (1,9) = 15.61, p = 0.003, η^2 = 0.63, a main effect of diagnostic status (F (1,9) = 12.91, p = 0.006, η^2 = 0.59), and an interaction (F (1,9) = 5.08, p = 0.05, η^2 = 0.36). Clinicians produced longer utterances to the older and the normal children, and the diagnostic status factor played a bigger role in the 12-year-old groups than in the 3-year-old groups.

Discussion

The existence of a therapeutic register is largely supported by our results. The clinicians were sensitive to both the age and the diagnostic status of their listeners. In terms of prosody, the clinicians slowed down for the impaired and the younger children. In terms of syntax, sentences were longer for the older and normal groups as compared to the younger and impaired groups.

Measures of verbosity (NTW, noun token, number of utterances) told a two-fold story. When the entire narrative was used as the unit of analysis, the four conditions were not differentiated. This was possibly due to the high SD values for these measures. The high variability indicated that these clinicians were not uniformly verbose or brief to a certain group of listeners. While some clinicians produced very short and simple stories to the 3-year-old impaired children, others went into detail to explain every picture. It may be the case that some clinicians used elaboration as a compensatory strategy while others regarded lengthiness as a hindrance for comprehension. On the other hand, when the verbosity measures were reproduced with utterance as the unit of analysis, the comparisons became highly significant. Taken together, the results seemed to indicate that while the clinicians provided an equally rich story to each group, they broke the input into smaller and manageable chunks of speech units for the impaired and the younger children to reduce the processing load and facilitate comprehension.

The diversity measures (NDW, noun type) demonstrated similar patterns of results as the verbosity measures. Clinicians exposed all groups of listeners to an

equally large set of vocabulary. In the meantime, they made the task manageable for the younger and impaired children by making fewer variations in each utterance. TTR—a measure of redundancy (Hayes and Ahrens, 1988), yielded a marginally significant effect for age. Clinicians were more repetitive to the younger children. The lack of main effect for diagnostic status for this measure could be a task-specific finding, because the storyteller had to maintain the flow of the story and was limited in her lexical choices.

We also found a few interactions in the analyses. All the interaction effects demonstrated a wider gap between the two older groups, indicating a higher degree of modification from 12-year-old normals to 12-year-old impaired. Despite the fact that the profiles of the normals were designed to be equally normal and those of the impaired were equally impaired, we still found this interaction. This pattern of results may have to do with the clinicians' attempts to simplify for 3-year-olds, whether impaired or normal, because of their young age or with clinicians' knowledge of the generally better prognoses for younger than older impaired children.

The existence of a therapeutic register was attested by the many main effects of diagnostic status in this simulated storytelling task. Though some of the effects were only marginally significant, this study constitutes a preliminary investigation of the therapeutic register with only a small subject pool and limited statistical power. Therefore, these effects were considered robust. Unlike previous findings, which showed that the presence of a child interactant was a necessity for the occurrence of motherese (Snow, 1972; Bohannon and Marquis, 1977), we found significant modifications of speech even in absence of a child listener in this group of clinicians. Diagnostic status played a role in addition to and independent of the well-known age factor. We believe that the role-play design, while ensuring experimental control at the expense of real clinician-client interaction, may have minimized the extent of therapeutic register usage, that is, we expect that findings from the present study would be generalizable to and magnified in real-life situations.

Effects concerning diagnostic status were mostly of the same magnitude as those related to the age variable, which suggests that the therapeutic register of interest could be regarded as a special form of motherese. However, we did find differences in two respects. Clinicians inserted longer pauses for the impaired children, but not for the younger ones. They produced more simple sentences for the 3-year-old groups, but not for the impaired groups. Further and more comprehensive studies are needed to determine the manner and extent to which motherese and the therapeutic register differ. The great individual differences in the present study also suggest that unlike motherese, which is argued to be biologically programmed in humans (Papoušek et al., 1987), the therapeutic register may take time and practice to acquire. Presently, we do not have data to support the prediction that there is a positive relationship between clinical experience and speech features tied to the therapeutic register. Evidence for this view would require studies that include sizable groups of both experienced and novice clinicians. The ultimate goal for studying the therapeutic register is to investigate the function of this register in clinical intervention and discover a package of features that constitute an effective intervention language. This would require longitudinal studies of clinician-client dyads and correlations between clinicians' speech features and clients' language growth.

Acknowledgements

The authors wish to thank Kanika So for obtaining the reliability measures and Xuejing Sun and Chi Xie for their help with programming in the acoustic analysis.

References

BOERSMA, P., 2001, *Praat: Doing Phonetics by Computer*, **4.0.11**, *http://www.praat.org/*.

BOHANNON, J. N. and MARQUIS, A. L., 1977, Children's control of adult speech. *Child Development*, **48**, 1002–1008.

CONTI-RAMSDEN, G., 1994, Language interaction with atypical language learners. In C. Gallaway and B. J. Richards (Eds) *Input and Interaction in Language Acquisition* (Cambridge: Cambridge University Press), pp. 183–196.

FERNALD, A. and SIMON, T., 1984, Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, **20**, 104–113.

GALLAWAY, C. and RICHARDS, B. J. (Eds), 1994, *Input and Interaction in Language Acquisition* (Cambridge: Cambridge University Press).

HAYES, D. P. and AHRENS, M. G., 1988, Vocabulary simplification for children: a special case of 'motherese'? *Journal of Child Language*, **15**, 395–410.

MAYER, M., 1974, *Two More Moral Tales* (New York: Four Winds Press).

PANAGOS, J. M., BOBKOFF, K. and SCOTT, C. M., 1986, Discourse analysis of language intervention. *Child Language Teaching and Therapy*, **2**, 211–229.

PAPOUŠEK, M., PAPOUŠEK, H. and HAEKEL, M., 1987, Didactic adjustments in fathers' and mothers' speech to their 3-month-old infants. *Journal of Psycholinguistic Research*, **16**, 491–516.

REUVERS, M. and HARGROVE, P. M., 1994, A profile of speech-language pathologists' prosody during language therapy. *Child Language Teaching and Therapy*, **10**, 139–152.

SHUTE, B. H., 1987, Vocal pitch in motherese. *Educational Psychology*, **7**, 187–205.

SNOW, C. E., 1972, Mother's speech to children learning language. *Child Development*, **43**, 549–565.

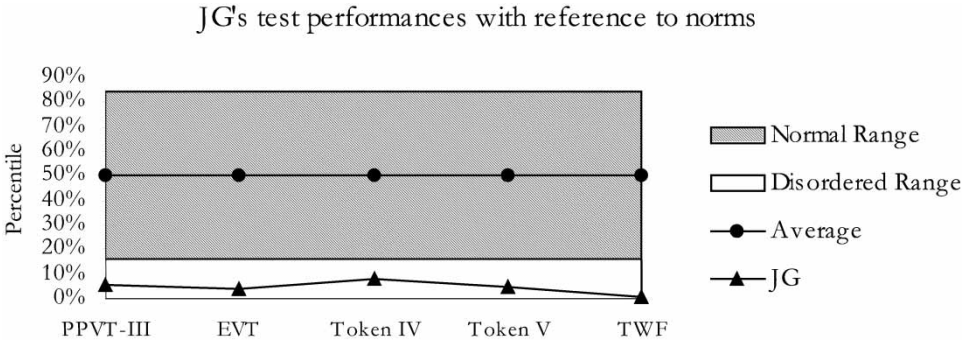
SNOW, C. E., 1994, Beginning from Baby Talk: 20 years of research on input in interaction. In C. Gallaway and B. J. Richards (Eds) *Input and Interaction in Language Acquisition* (Cambridge: Cambridge University Press), pp. 3–12.

SNOW, C. E. and FERGUSON, C. A. (Eds), 1977, *Talking to Children: Language Input and Acquisition* (Cambridge: Cambridge University Press).

STERN, D. N., SPIEKER, S. and BARNETT, R. K., 1983, The prosody of maternal speech: infant age and context related changes. *Journal of Child Language*, **10**, 1–15.

Appendix. Language profile of an impaired adolescent

Name: JG Age: 12;1 Gender: F



PPVT-III: The Peabody Picture Vocabulary Test (3rd Edition) measures single word vocabulary comprehension.

EVT: The Expressive Vocabulary Test assesses vocabulary naming and the ability to retrieve words from memory.

Token IV and *Token V*: *The Token Tests* assess auditory comprehension and temporal and spatial concepts.

TWF: The Test of Word Finding measures single word retrieval.

Sentence imitation: JG correctly repeated only one sentence (six words in length). She retained the gist of sentences up to about 11 words in length, but made errors in surface structure. She omitted major segments of longer sentences.

Word sequence: JG correctly repeated sequences of three unrelated words. She did not correctly repeat sequences of four unrelated words.