Paul Corthals

Effects of speaker gender and child age on the prosody of parentese : cross-linguistic evidence

Logopedia Silesiana 3, 19-30

2014

Paul Corthals
Ghent University / University College, Belgium

Effects of speaker gender and child age on the prosody of parentese: cross-linguistic evidence

Abstract: „Spieszczenia" to wyrażenia językowe, jakimi posługują się dorośli, mówiąc do dzieci. Ich funkcją jest regulowanie pobudzenia i komunikowanie emocji. Ich suprasegmentalna prozodia ułatwia naukę języka. Ta „pouczająca prozodia" jest porównywana w języku angielskim i holenderskim oraz pomiędzy mową męską i żeńską, adresowaną do dzieci w wieku 0,5–2,4 lat. Zostały szczegółowo określone różnice w prozodii pomiędzy próbkami mowy kierowanej do dziecka i mowy kierowanej do dorosłych. Wykonano analizę ANCOVA, traktując parametry prozodii jako zmienne zależne, płeć dorosłych oraz język – jako stałe, a wiek dziecka – jako zmienną towarzyszącą. Kobiety bardziej uwypuklały intonację i modulowały ton głosu adekwatnie do wieku dziecka. Ton głosu był podniesiony dużo bardziej w grupie kobiet mówiących w języku amerykańskim. Szybkość mówienia była istotnie niższa w języku flamandzkim, holenderskim. Hipoteza zakłada, że ton głosu jest używany jako parajęzykowa cecha adresowana do młodszych niemowląt, a niższe tempo mówienia – do starszych dzieci, które zaczynają rozumieć komunikaty werbalne.

Key words: prozodia, intonacja, suprasegmentalna prozodia, badania mowy dziecka

Introduction

Parentese is the speech-language register adults switch to when talking to children. It uses a limited vocabulary, short repeated utterances, and referents that are concrete and present. Acoustic-phonetic analyses show hyper-articulated vowels, slower speech and articulation rate, raised voice pitch, exaggerated intonation, and pre-boundary vowel lengthening. Parentese probably regulates infant arousal and attention and communicates affect[1], but it is also thought to facilitate language learning because it's linguistic complexity is tuned to children's language develop-

[1] A. Fernald: *Intonation and communicative intent in mothers' speech to infants: is the melody the message?* "Child Development" 1989, Vol. 60, p. 1497–1510.

ment stages and it's segment-marking prosody can act as a disambiguating factor. This study concentrates on the suprasegmentals of parentese.

The role ascribed to the phonetical aspects of ambient language input in the process of language acquisition has changed with successive theoretical positions. According to Skinnerian reinforcement-based interpretations, the importance of ambient language and speech was marginal. In the Chomskyan nativist interpretation, speech and language input triggersthe selective maturation of certain innate grammatical and phonetic abilities that are relevant in the child's own mother tongue and, at the same time, the disuse and decline of non-native abilities. In 1989, Snow and Ferguson redirected the line of thought by terming the distinctive speech and language style demonstrated by caregivers interacting with young children as 'motherese' (the label 'parentese' came later). In recent views on language acquisition[2] parentese is seen as the catalyst element that helps young infants to detect phonotactic regularities, prosodic patterns and prototypes of phonetic units, altering innate perceptual abilities, and allowing them to bootstrap major constituent boundaries in the first year of life before they can speak.

Infants track distributional information about phonemes and syllables[3]. For example, syllables that frequently co-occur are likely to be part of the same word, whereas syllables that co-occur only rarely most of the time span a word boundary. These co-occurrence statistics are known as transitional probabilities.

Apart from transitional probabilities, also prosody helps in the process of parsing speech. Prosodic adaptations in child-directed speech include pitch (such as a declination reset at the end of clauses), rhythm (such as pausing and pre-boundary vowel lengthening). Relative loudness, timing and intonation inchild-directed speech are seen as acoustic cues for even young infants to grammatical structure, utterance segmentation and word identification. There seem to be universalas well as language-specific features. D.L. Grieser & P.K. Kuhl[4] found similar features in infant-directed speech in a tonal language (Mandarin Chinese) and non-tonal languages (English and German), such as a higher fundamental frequency, a larger range of intonation maneuvers as well as a slower speech rate. C. Fisher & H. Tokura[5] studied American English and Japanese parentese and found systematic lengthening and exaggerated pitch changes in utterance-final vowels in both languages as well as language-specific syllable duration and pitch effects associated with phrase

---

[2] P.K. Kuhl, B.T. Conboy, S. Coffey-Corina, D. Padden, M. Rivera-Gaxiola, T. Nelson: *Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e).* "Philosophical Transactions of the Royal Society B" 2008, Vol. 363, p. 979–1000.

[3] E.D. Thiessen, J.R. Saffran: *When Cues Collide: Use of Stress and Statistical Cues to Word Boundaries by 7- to 9- Month-Old Infants.* "Developmental Psychology" 2003, Vol. 39, p. 706–716.

[4] D.L. Grieser, P.K. Kuhl: *Maternal speech to infants in a tonal language: support for universal prosodic features in motherese.* "Developmental Psychology" 1988, Vol. 24, p. 14–20.

[5] C. Fisher, H. Tokura: *Acoustic cues to grammatical structure in infant-directed speech: cross-linguistic evidence.* "Child Development" 1996, Vol. 67, p. 3192–3218.

boundaries within utterances. R. Shi, J.L. Morgan & P. Allopena[6] compared distributional, phonological and prosodic features of both lexical and functional word categories in two Mandarin and two Turkish samples of infant-directed speech. The prosodic features involved were vowel or syllable duration, relative amplitude and pitch change. In each language, lexical and functional items clearly differed, but none of the measures could predict grammatical category membership on its own. However, in each language, constellations of cues were sufficient to assign words to grammatical categories. E. Payne, B. Post, L. Astruc, P. Prieto & M. Vanrell[7] found a greater uniformity of overall syllable duration combined with specific phrase-final prosodic lengthening in a study of English and Catalan child-directed speech. In English, there also was durational marking of the nuclear accented syllable.

The evidence of "didactic prosody" in child-directed speech is completed by studies demonstrating that language learning is indeed facilitated by it. A.D. Endress & M.D Hauser[8] showed that adult English speakers are able to segment utterances and recognize words from fluent speech in foreign languages they had no prior exposure to, using prosodic cues only. There is evidence that infants and children can do the same. T. Nazzi, T. Kemler, D.G. Nelson, P.W. Jusczyk & A.M. Jusczyk[9] studied the headturn reactions of six-months-olds to spoken passages containing to prosodically well-formed and ill-formed sequences of the same words. The infants recognized the well-formed sequence better, which is evidence for early prosodic bootstrapping strategies. A. Seidl[10] repeated the T. Nazzi[11] et al. experiment, but neutralized one or more acoustic correlates of clausal boundaries in the stimuli, i.e. pitch, pause or vowel duration. Her conclusion was that pitch is necessary but not sufficient as a cue for six-month-olds to segment clauses (infants relied on the combination of pitch and pause or pitch and vowel length) and that infants of this age already apply a certain weighting to prosodic cues. M. Soderstrom, A. Seidl, D.G. Kemler Nelson & P.W. Jusczyk[12] provide evidence

[6] R. Shi, J.L. Morgan & P. Allopena: *Phonological and acoustic bases for earliest grammatical category assignment: a cross-linguistic perspective*. "Journal of Child Language" 1998, Vol. 25, p. 169–201.

[7] E. Payne, B. Post, L. Astruc, P. Prieto, M. Vanrell: *A cross-linguistic study of prosodic lengthening in child-directed speech*. Chicago, Speech Prosody 2010.

[8] A.D. Endress, M.D Hauser: *Word segmentation with universal prosodic cues*. Cognitive Psychology 2010, Vol. 61, p. 177–199.

[9] T. Nazzi, T. Kemler, D.G. Nelson, P.W. Jusczyk, A.M. Jusczyk: *Six-month-olds detection of clauses embedded in continuous speech: effects of prosodic well-formedness*. "Infancy" 2000, Vol. 1, p. 123–147.

[10] A. Seidl: *Infants' use and weighting of prosodic cues in clause segmentation*. "Journal of Memory and Language" 2007, Vol. 57, p. 24–48.

[11] T. Nazzi, T. Kemler, D.G. Nelson, P.W. Jusczyk, A.M. Jusczyk: *Six-month-olds detection…*

[12] M. Soderstrom, A. Seidl, D.G. Kemler Nelson, P.W. Jusczyk: *The prosodic bootstrapping of phrases: evidence from prelinguistic infants*. "Journal of Memory and Language" 2003, Vol. 49, p. 249–267.

that infants as young as 6 months of age are sensitive to syntax-related prosodic cues that signal units smaller than a clause. P.W. Jusczyk & D.M. Houston[13] and V. Kooijman, P. Hagoort, A. Cutler[14] demonstrated that 7.5- and 10-month-olds rely on typical prosodic word patterns to segment words in running speech.

Emanating from these studies is a picture of development in receptive prosodic competence that is proceeding from early infant sensitivity to the prosodic features in ambient language[15] towards the ability to exploitprosody for parsing-parentese into larger clause-level units at first and into phrase- and word-level units later on.

Given the relevance of prosody for early language processing and the evolution infants and children go through, it is interesting to explore language-related and gender-related variations among parentese speakers and child age-related evolution in the prosody of parentese. A. Warren-Leubecker & J.N. Bohannon[16] recorded dyadic sessions of mothers and fathers interacting with children of two and five years of age. They found gender differences and an interaction between sex of the speaker and the age of the listener for modal voice pitch elevation and voice pitch range (mothers increased intonation ranges more when addressing younger children; fathers did not differentiate voice pitch between 5-year-olds and adult listeners). A. Fernald, T. Taeschner, J. Dunn, M. Papousek, B. de Boysson-Bardies & I. Fukui[17] also found differences in the range of the fundamental frequency of the voice in male and female infant-directed speech. In a comparison of several language groups, they noted the largest prosodic modifications in American English parentese. In general, the prosodic exaggeration seems more acute in American English than in other languages and even other dialects of English[18].

B. Shute & K. Wheldall[19] compared voice pitch and speech rate in male and female parentese speakers. For conversational speech modes, they found larger standard deviations of voice pitch in mothers and a higher mean fundamental in

---

[13] P.W. Jusczyk, D.M. Houston: *The beginnings of word segmentation in English-learning infants.* "Cognitive Psychology" 1999, Vol. 39, p. 159–207.

[14] V. Kooijman, P. Hagoort, A. Cutler: *Prosodic structure in early word segmentation: ERP evidence from Dutch ten-month-olds.* "Infancy" 2009, Vol. 14, p. 591–612.

[15] B. Höhle, R. Bijeljac-Babic, B. Herolda, J. Weissenborn& T. Nazzi: *Language specific prosodic preferences during the first half year of life: evidence from German and French infants.* "Infant Behavior & Development" 2009, Vol. 32, p. 262–274.

[16] A. Warren-Leubecker & J.N. Bohannon: *Intonation patterns in child-directed speech: mother-father differences.* "Child Development" 1984, Vol. 55, p. 1379–1385.

[17] A. Fernald, T. Taeschner, J. Dunn, M. Papousek, B. de Boysson-Bardies, I. Fukui: *A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants.* "Journal of Child Language" 1989, Vol. 16, p. 477–501.

[18] M. Soderstrom: *Beyond babytalk: re-evaluating the nature and content of speech input to preverbal infants.* "Developmental Review" 2007, Vol. 27, p. 501–532.

[19] B. Shute, K. Wheldall: *Fundamental frequency and temporal modifications in the speech of british fathers to their Children.* "Educational Psychology" 1999, Vol. 19, p. 221–233,

fathers. H.-M. Liu, F.-M. Tsao & P.K. Kuhl[20] recorded Mandarin parentese produced by mothers producing twelve preselected bisyllabic nouns in a semi-structured situation. The same mother-child dyads were recorded at two points in time (child ages 0;7 and 5;0).When addressing their preverbal infants, mothers spoke more slowly, raised their pitch higher and used a wider F_0 range than in interactions with five-year-olds.

It is clear that parentese is not an undifferentiated or rigid whole. Its content en form depend on many factors, such as native language, speaker gender, and child age. Moreover, the initial preference infants have for typical child-directed speech seems to fade, suggesting that children access different aspects of speech-language input at different stages in their development[21].

The aim of this study is to compare the "didactic prosody" of child-directed speech in English and Dutch, and to juxtaposemale and female prosodic styles in parentese addressed to children between 5 and 28 months old.

Method

Participants

There were 19 female participants (mean age 35 yrs., standard deviation 10.2 yrs.) and 12 male participants (mean age 36 yrs., standard deviation 11.4 yrs.). Participants were recruited Virginia, USA (9 females and 5 males, all native speakers of American English, mean age 32 yrs., standard deviation 4.6 yrs.) and in Flanders, Belgium (10 females and 7 males, all native speakers of Flemish Dutch, mean age 38 yrs., standard deviation 13.1 yrs.). They agreed to being audio-recorded during a dyadic session with a child and during a short conversation with one of the investigators.

The children were the sons and daughters of the adult participants themselves or their close family. All children were typically developing (between 5 and 28 months, mean age 15 months, standard deviation 7.1 months). The American children (4 girls and 10 boys) were 11 months old on average (standard deviation 5.6 months). The Belgian children (5 girls and 12 boys) had a mean age of 19 months (standard deviation 5.9 months).This mean age difference of 8 months between the American and the Belgian children is statistically significant ($T = 3.977$, $p < 0.001$). The children who interacted with male participants did not differ in age from those

[20] H.-M. Liu, F.-M. Tsao, P.K. Kuhl: *Age-related changes in acoustic modifications of Mandarin Maternal speech to preverbal infants and five-year-old children: a longitudinal study.* "Journal of Child Language" 2009, Vol. 26, p. 909–922.

[21] M. Soderstrom: *Beyond babytalk*…

interacting with female participants (in fact, the same child often acted twice as a conversation partner in each gender subgroup).

Recordings

Digital recordings were made in each subject's home, using a laptop and a microphone. The same apparatus was used for both the adult-directed (AD) and child-directed (CD) samples. For AD speech samples, participants were encouraged to respond to unscripted questions. They were then asked to verbally interact with their child using a book or toy for the CD recordings. No directives were given as to the content or style of speaking.

Analysis

Praat software[22] was used to identify relevant fragments in the recordings. A fragment was considered relevant if there was only one voice in it (the adult participant's voice) and if the words spoken were recognizable. Fragments were delimitated based on visual inspection of time-amplitude diagrams and spectrograms and on auditory verification during replay. This selection resulted in a series of analyzable stretches per sample, amounting to approximately two minutes and containing 200 up to 400 words.

Praat software was also used to retrieve parameters for speech and articulation rate (words and syllables per unit of time), voice pitch (median of the voice fundamental, extent and speed of the intonation maneuvers) and voice intensity (extent and speed of the voice stress maneuvers). As a first step, text tiers were added to the sound files to note the number of words and syllables within each relevant fragment. A script was programmed to determine the duration of relevant fragments and to retrieve the word and syllable counts, in order to calculate speech rate and articulation rate. Voice pitch was examined by means of another script extracting the median and the interquartile range of the fundamental frequency from all voiced parts in the relevant fragments. Also, the sum of the absolute value of all F_0 changes between the 25th and 75th percentile was calculated, cumulated from the start to the end, and divided by the total duration of the relevant utterances. This index, expressed in Hz/sec, reflects both extent and speed of the intonation maneuvers. Similar indices for voice intensity were determined, i.e. the interquartile range as an index for modulation depth and the sum of all intensity variations within the interquartile range divided by total duration as a dynamic index for extent and speed

[22] P. Boersma, D. Weenink: *Institute of Phonetics*. Amsterdam 2004. PRAAT [computer software].

voice stress maneuvers. Values out of the interquartile ranges were excluded in order to prevent outliers and artifacts from influencing the results. The absolute value of median voice intensity was not used in the analysis, as the mouth-to-microphone distance could not always be kept identical in both recording sessions.

Statistical processing

The SPSS19 software package was used for statistical processing. Differences were calculated per speaker between the CD and the AD sample (child-directed values minus adult-directed values) for the rate, pitch and intensity parameters. One-Sample-Kolmogorov-Smirnov Tests were used to verify the distribution of these rate, pitch and intensity contrasts. A series of gender (2) language (2) ANCOVA analyses were performed, treating the prosodic AD-CD contrastsas dependent variables, adult gender and language as fixed factors, and child age as a covariate. Levene's test was done to check the homogeneity of variance assumption and the interaction terms with the covariate were verified to check for the homogeneity of regression slopes. For the significant effects found in the ANCOVA analyses, scatter plots and trend lines were studied to check the evolution over time (i.e. with growing child age).

Results

The distributions of the rate, pitch and intensity contrasts were normal. The assumptions of homogeneity of variance and of regression slopesweremet. In the ANCOVA outcomes, there were no significant gender effects except for the pitch modulation AD-CD contrast (interquartile range of the fundamental frequency, $F(1.25)= 16.753$, $p <0.001$, partial eta squared = 0.401), and no significant language effects except for the speech rate AD-CD contrast (number of words per unit of time, $F(1.29) = 4.843$, $p = 0.036$, partial eta squared = 0.143). There was a significant gender-language interaction for the voice pitch AD-CD contrast (median of fundamental frequency, $F(1.25)= 7.934$, $p =0.009$, partial eta squared = 0.241).

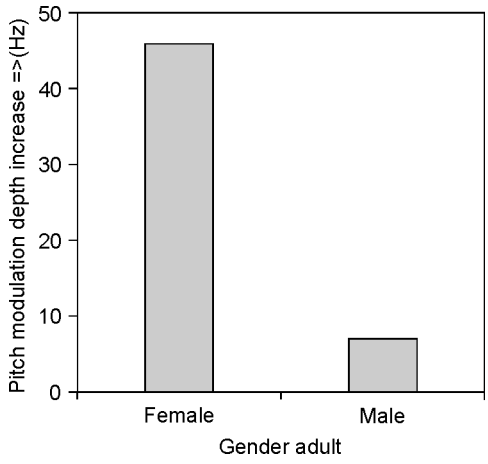

FIGURE 1. Gender effect in expanding the range of intonation maneuvers when a speaker switches to child-directed speech

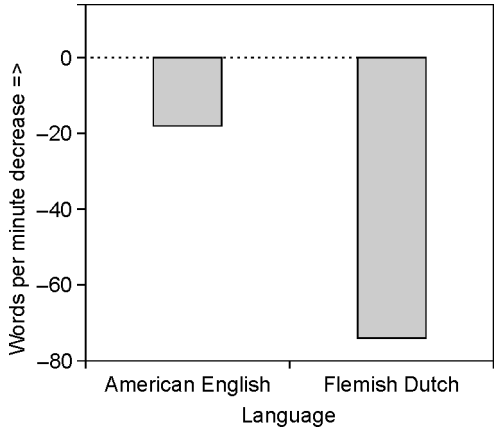

FIGURE 2. Language effect in slowing down speech rate when a speaker switches to child-directed speech

The gender effect in pitch modulation depth is illustrated in figure 1. Female parentese speakers expand the range for intonation maneuvers much more than male speakers do: the mean AD-CD contrast for interquartile range of the fundamental frequency was 47 Hz (standard deviation 30.2 Hz) in female samples versus 6 Hz (standard deviation 21.3 Hz) in male samples. The language effect in slowing down speech rate is illustrated in figure 2. Speech rate was lowered more in Flemish Dutch parentese: the mean AD-CD contrast for the number of words per unit of time was minus 72 (standard deviation 38 words) in Flemish Dutch samples versus minus 16 (standard deviation 58 words) in American English samples. The language-gender interaction for voice pitch increment is illustrated in figure 3. Voice pitch raise is clearly more pronounced in American female parentese speakers (mean AD-CD contrast forvoice pitch 53 Hz, standard deviation 51.7 Hz). Figures 4 and 5 show the AD-CD contrasts of median voice pitch and speech rate as a function of the child's age for, both adult gender groups (figure 4) and for male speakers separately (figure 5). For the youngest children, voice pitch is raised (particularly by

female speakers), whereas the trend line for speech rate points to a status quo. For the oldest children, voice pitch approaches the levels typical of adult conversation, whereas speech rate shows a decreasing trend.

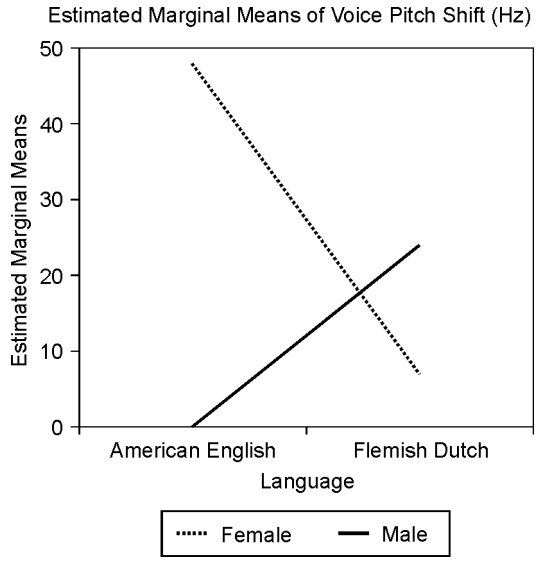

Figure 3. Language-gender interaction for voice pitch heightening when a speaker switches to child-directed speech

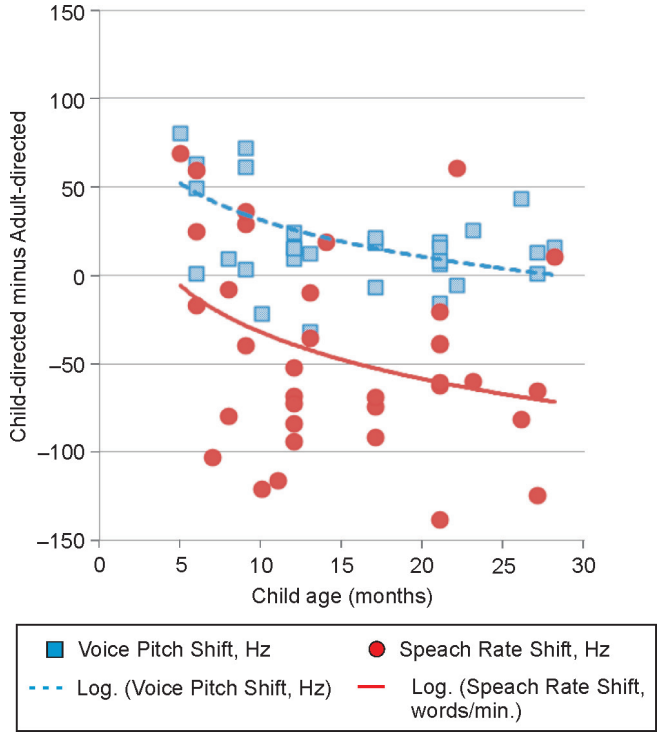

Figure 4. Adaptations in habitual voice pitch and speech rate in parentese as a function of the child's age (both adult gender groups)

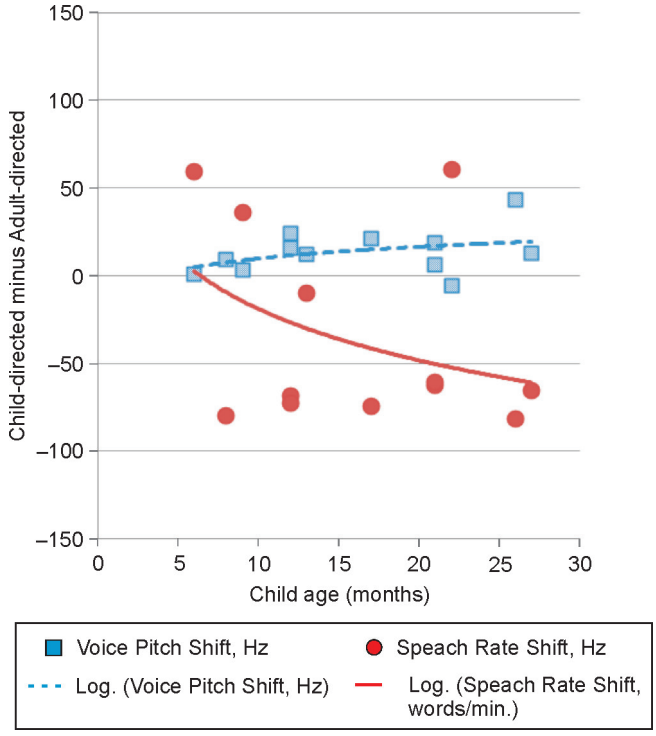

Figure 5. Adaptations in habitual voice pitch and speech rate in parentese as a function of the child's age (male speakers only)

Discussion

Effects of gender and child age

In both American English and Flemish Dutch parentese, prosody had gender-specific features in that pitch modulation depth was significantly larger in female speakers. A. Warren-Leubecker & J.N. Bohannon[23], A. Fernald[24], and several other researchers[25] have also found more prominent pitch excursions in female speakers.

As illustrated in figures 4 and 5, median voice pitch seems to be tuned to the child's age in female parentese only. The trend line for voice pitch heightening in figure 4 starts at a raise of about 50 Hz and falls back to zero for the oldest children in the study. This tendency remains unchanged whenoutcomes are plotted

[23] A. Warren-Leubecker, J.N. Bohannon: *Intonation patterns in child-directed speech…*
[24] A. Fernald: *Intonation and communicative intent in mothers' speech…*
[25] M. Soderstrom: *Beyond babytalk…*

for female speakers a separate group, but it is absent in male parentese samples (figure 5). A. Warren-Leubecker & J.N. Bohannon[26], who compared child-directed speech addressed to two- and five-tear-olds to adult-directed speech, noted that men raised their pitch and pitch ranges when speaking to the younger children, but did not differentiate between 5-year-olds and adult listeners. They hypothesized that men may revert to the more monotonic speech patterns of adult-directed speech in an attempt to avoid a stereotypically feminine speech style. H.-M. Liu et al.[27] also found an attunement of voice pitch elevation to the child's age in the mother-child dyads with the youngest children in their study, but this apparent female parentese sub-style remains a less well documented phenomenon.

Effects of language and child age

Prosodic exaggeration may be more prominent in American English than in other languages[28], but the language-related effects in the present study should be interpreted cautiously. Although the age of the child was treated as a covariate, the main language effect (speech rate was lowered more in Flemish Dutch, figure 2), and the language-gender interaction (larger voice pitch shift in female speakers of American English, figure 3) should be seen in the light of the fact that the American children were younger than the Belgian children. The 8 month age difference may have had consequences for the nature of adult-child interactions. One would expect an evolution in parentese from conveying mostly affective and social-regulatory intentions in the beginning, to more and more informative contents later on as the child's development progresses. If communicative intentions behind parentese indeed change as a function of the child's age in this manner, then voice pitch is an obvious paralinguistic feature to regulate infant arousal and attention and to communicate affect when addressing young infants, whereas a lower speech rate is needed for messages addressed to older children, who start to comprehend them as verbal signals.

This intention-based interpretation of parentese prosody is corroborated by our findings on pitch-age tuning and rate-age tuning. The largest voice pitch raises were found in female speakers interacting with the youngest children, which is compatible with the primacy of affective and social-regulatory communicative intentions. These adaptations of median voice pitch gradually decline with growing age of the child, to be replaced by adaptations in speech rate, in line withthe shift towards conveying information that is hypothesized here. The trend line for the

[26] A. Warren-Leubecker, J.N. Bohannon: *Intonation patterns in child-directed speech…*

[27] H.-M. Liu, F.-M. Tsao, P.K. Kuhl: *Age-related changes in acoustic modifications…*

[28] A. Fernald, T. Taeschner, J. Dunn, M. Papousek, B. de Boysson-Bardies, I. Fukui: *A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants.* "Journal of Child Language" 1989, Vol. 16, p. 477–501; M. Soderstrom: *Beyond babytalk…*

amount of speech rate lowering in figure 4 starts at zero and ends at about minus 60 words per minute for speakers addressing the oldest children in the study. This tendency is also present in male parentese samples (figure 5).

It is well known that slower speaking rates result in a larger vowel space relative to that measured for faster rates[29]. Thus, slowing down the speech rate underpins speech clarity and intelligibility, which again is in line with the explanation that pitch and rate modifications in parentese may be related to the development of the child and the type of adult-child interactions across time.

Under this hypothesis, there is no genuine main language effect for speech rate (figure 2), and no authentic language-gender interaction for voice pitch raise (figure 3), as both findings are conceivedas side effects of the change in communicative intentions over time. Further research is needed to resolve this issue by comparing prosody and contents of parentese addressed to balanced age groups in both languages.

---

[29] M. Fourakis: *Tempo, stress, and vowel reduction in American English*. "Journal of the Acoustical Society of America" 1991, Vol. 90, p. 1816–1827.