

Mothers Are Less Efficient in Employing Prosodic Disambiguation in Child-Directed Speech than Non-Mothers: Is There a Trade-Off Between Affective and Linguistic Prosody?

Sonja Schaeffler[†] & Vera Kempe[‡]

[†]SSRC, Queen Margaret University, Edinburgh, UK, [‡]Dept of Psychology, Stirling University, UK
sschaeffler@qmu.ac.uk, vk2@stir.ac.uk

ABSTRACT

This study examines prosodic disambiguation in child-directed (CD) speech. Twenty-four mothers addressed syntactically ambiguous sentences to their 2;0 to 3;8 year old child and to an adult confederate. Twenty-four non-mothers addressed an imaginary toddler and an imaginary adult. We found that only mothers increased pitch and produced the CD-typical pitch excursions when addressing their children. In contrast, non-mothers, but not mothers, used prosodic disambiguation in CD speech, which was corroborated by a forced choice test in which 48 listeners judged the intended meaning of each sentence. The results suggest that if speakers express genuine positive affect, they tend to emphasise affective prosody at the expense of linguistic prosody. In the case of CD speech, this communication strategy may be more effective as it serves to elicit the child's attention.

1. INTRODUCTION

[12] demonstrated that when the referential context supports various interpretations of a syntactically ambiguous sentence, speakers are able to disambiguate using prosody if they are aware of the contextual ambiguity. For ambiguous sentences such as *Tap the frog with the flower* the most prominent prosodic cues were the duration of the pause before the prepositional phrase (PP), and the presence or absence of a pitch accent on the preposition 'with'. These prosodic cues indicate whether the prepositional phrase 'with the flower' is attached to the verb phrase 'tap' or the transitive object noun phrase 'the frog'. The former would correspond to an interpretation of the PP as an instrument; the latter would render it a modifier.

We use the paradigm developed by [12] to examine prosodic disambiguation in CD speech. According to the Prosodic Bootstrapping Hypothesis [9], CDS contains prosodic cues to syntax which may help children to acquire the structure of language. Under this hypothesis, we expect mothers to

provide clearer prosodic cues for disambiguation in CD compared to AD speech.

However, it has been argued that the primary function of CD is not didactic, but to influence the child's attention and arousal [3]. This kind of emotion manipulation may be a by-product of the affective prosody of CD speech [11] that arises from the emotional bond with the child [10]. The interesting question is whether affective prosody associated with positive affect expression results in clearer prosodic cues to syntactic structure, particularly in situations of ambiguity.

In this study, mothers and non-mothers completed a referential communication task containing syntactically ambiguous sentences addressed to their child vs. an imaginary child. Non-mothers addressing an imaginary child were assumed to lack the emotional bond with a child that is responsible for affective prosody. Thus, the comparison between mothers and non-mothers allowed us to study the effect of affect expression on prosodic cues to syntactic structure.

2. EXPERIMENT 1: Mothers

2.1. Method

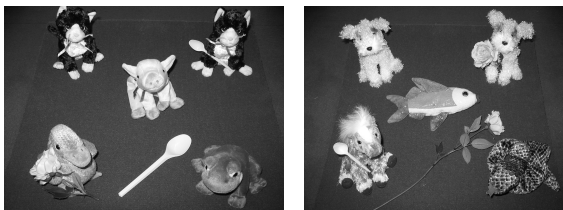
Participants: Twenty-four mothers, aged 23-46 years, and their children, aged 2;0-3;8 years, completed the referential communication task. Twenty-four undergraduate students (8 men), aged 19-50 years performed a forced-choice comprehension task.

Materials: Two arrays of toy objects provided the referential contexts (Figure 1). Each array contained two exemplars of the referent for the transitive object (a cat vs. a dog), one of which was holding a minor object (a spoon vs. a flower), as well as a larger version of this object which could serve as an instrument for touching. In addition, the arrays contained two other animals, one of which was also holding a minor object. This allowed us to create sentences with PP attachments that were am-

biguous syntactically as well as with respect to the referential context as in (1a) and (1b), and sentences with PPs that were syntactically ambiguous but were readily disambiguated by the referential context as in (2a-d). We also added three filler sentences per array.

- (1a) Touch the cat with the spoon.
- (1b) Touch the dog with the flower.
- (2a) Touch the fish with the flower (instrument)
- (2b) Touch the duck with the flower (modifier)
- (2c) Touch the frog with the spoon (instrument)
- (2d) Touch the horse with the spoon (modifier)

Figure 1: Object arrays used in the experiments.



Using the toy objects, we created digital colour photographs of the actions speakers had to describe using the sentences. For the instrument interpretation, the photographs showed a hand holding the object (e.g. the spoon) touching the designated toy (i.e. the cat without the little spoon). For the modifier interpretation, the photographs showed a hand touching the designated toy (i.e. the cat holding the little spoon). The intended actions for sentences with unambiguous referential context and filler sentences were depicted accordingly.

The photographs, with the corresponding sentence printed underneath, were assembled into two booklets which differed only in the intended action of the ambiguous sentences. Thus, if in Booklet 1 a critical sentence was presented with an action corresponding to an instrument interpretation, then in Booklet 2 it was presented with a modifier interpretation, and vice versa. Presentation of booklets was counterbalanced. We also created digital photographs containing just the two possible actions for each sentence, without the remaining referential objects. For example, for sentence (1a), one photograph showed a cat being touched by a spoon (instrument interpretation), and the other photograph showed a cat holding a little spoon being touched by a hand (modifier interpretation). These photographs were presented next to each other in the forced-choice task using E-prime.

Procedure and Measurements: The mothers, with their children seated on their lap in front of

the array of toys, were shown the booklets, and asked to use the provided sentences to instruct the child to perform the depicted actions. In the adult-directed (AD) condition, they were asked to do the same addressing an adult confederate. Mothers were instructed to carefully study each picture before producing the accompanying sentence. Order of CD and AD conditions was counterbalanced. The mothers' speech was audio-recorded.

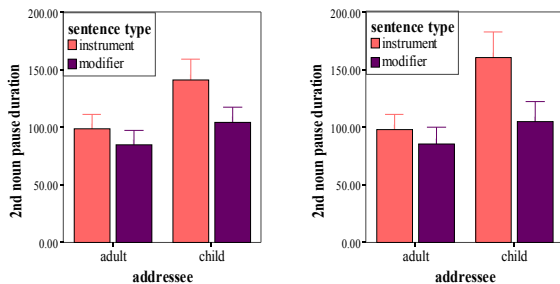
The recordings of the 192 critical sentences of the type 'Touch the N1 with the N2.' were then combined with the action photographs, and presented to 24 other participants in a forced-choice task asking them to decide which picture corresponded to the intended meaning of the sentence. The sentences were also submitted to acoustic analyses measuring the duration of verb pause and noun pause, and F0 on the steady-state portion of the vowels in N1, 'with', and N2 (in Hz). For the bi-syllabic 'flower', F0 was measured for the entire triphthong. Pauses were identified through visual inspection of the spectrogram and the acoustic signal. F0 values were obtained using PRAAT algorithms [2].

2.2. Results and Discussion

Unless detailed otherwise, all dependent variables were analysed using a 2 (Context Ambiguity: ambiguous vs. unambiguous) x 2 (Addressee: adult vs. child) x 2 (Sentence Type: instrument vs. modifier) ANOVA.

Prosodic features: For the verb pauses, the ANOVA revealed a main effect of Addressee, $F(1,23)=5.2$, $p<.05$, indicating that pause durations increased in the CD condition, most likely due to an overall decrease in speech rate. For the noun pauses, the ANOVA revealed a main effect of Addressee, $F(1,23)=6.4$, $p<.05$, due to longer pause durations in the CD condition. The effects of Context Ambiguity, $F(1,23)=2.7$, $p=.1$, and the effect of Sentence Type, $F(1,23)=2.9$, $p=.1$, fell short of significance suggesting that noun pauses tended to be shorter when the context was ambiguous, and when the mothers intended a modifier interpretation. Thus, mothers showed the expected tendency towards shorter pauses before the PP in the modifier condition. However, the lack of an interaction between Addressee and Sentence Type suggests that prosodic disambiguation was not more pronounced when speaking to children (see Figure 2, left panel).

Figure 2: 2nd (noun) pause durations for the instrument and modifier interpretations in mothers (left panel) and non-mothers (right panel).



F0 analyses sought to determine whether there were pitch accents on the preposition ‘with’ which could serve as a prosodic cue for the instrument interpretation, and also to what extent mothers tended to raise pitch in the CD condition. The ANOVA, which additionally contained the factor Position (N1, ‘with’, N2) revealed a main effect of Addressee, $F(1,23)=64.1$, $p<.001$ indicating that, as expected, F0 was higher in the CD condition. There was a main effect of Position, $F(2,46)=5.25$, $p<.05$. Post-hoc tests using Fisher’s LSD revealed a difference in F0 between the preposition ‘with’ (217 Hz) and N2 (247 Hz). We also found an interaction between Position and Context Ambiguity, $F(2,46)=4.2$, $p<.05$, and between Position and Addressee, $F(2,46)=4.5$, $p<.05$. Thus, mothers increased F0 when addressing their children, particularly on N2 at the end of the sentence, a pattern that would be expected in CD speech [4]. Interestingly, this intonation pattern was more pronounced in ambiguous sentences suggesting that, rather than placing a pitch accent on the preposition when the context was ambiguous, as the speakers in [12] did, the mothers amplified the CD-typical final pitch-excursion pattern. Thus, mothers may have been well aware of the ambiguity of some contexts, but did not use effective prosodic means to disambiguate, opting instead for more ‘child-directedness’.

Forced-choice comprehension task: The ANOVA on the percent of instrument choices revealed a main effect of Sentence Type, $F(1,23)=5.7$, $p<.05$, indicating that there were more instrument choices in instrument sentences (56%) than in modifier sentences (49%). There was also a main effect of Context Ambiguity, $F(1,23)=10.2$, $p<.01$, indicating that instrument choices were more frequent in unambiguous (54%) than ambiguous (51%) sentences. Finally, the main effect of Addressee, $F(1,23)=4.4$, $p<.05$, was due to more instrument choices in the CD condition (54.8%) than in the

AD condition (49%) indicating that listeners were sensitive to the general increase in noun pause duration in CD speech. Note that comprehension of the intended interpretation was quite poor with performance generally around chance (50%). One-sample t-tests against chance showed that only in the unambiguous AD and CD instrument sentences, and in the ambiguous AD modifier sentences were listeners able to reliably identify the intended interpretation. In all other conditions performance was indeed at chance. This shows that the prosodic cues the mothers provided were generally not sufficiently clear to the listeners to enable them to identify the intended interpretation. As in the pause data, the lack of an interaction between Addressee and Sentence Type suggests that there was no greater clarity in the disambiguating prosodic cues in the CD condition.

3. EXPERIMENT 2: Non-mothers

3.1. Method

Participants: Twenty-four non-mothers, aged 21-42 years, completed the referential communication task. Twenty-four undergraduate students (7 men), aged 18-42 years performed the forced-choice comprehension task.

Materials: Materials were identical to Experiment 1.

Procedure and Measurements: Procedure and measurements were identical to Experiment 1 except that non-mothers were told that their instructions would be played back to listeners to test a ‘game’. They were asked to address an imaginary adult as if speaking to an acquaintance, and to address an imaginary child as if speaking to a 2-3-year old. Order of imaginary addressee (adult vs. child) was counterbalanced across participants.

3.2. Results and Discussion

Prosodic features: For the verb pauses, the ANOVA showed no significant effects. For the noun pauses, we found a main effect of Addressee, $F(1,23)=8.6$, $p<.01$, due to longer pause durations in the CD condition. The main effect of Sentence Type, $F(1,23)=6.3$, $p<.05$, was specified by an interaction between Sentence Type and Addressee, $F(1,23)=5.7$, $p<.05$: the difference in noun pause duration between instrument and modifier sentences was significant in the CD condition, $F(1,23)=8.3$, $p<.01$, but not in the AD condition, $F<1$ (see Figure 2, right panel). Interestingly, there

was no difference between ambiguous and unambiguous contexts suggesting that non-mothers did not employ audience design [7].

For F0, the effect of Position fell short of significance, $F(2,46)=3.4$, $p=.06$. Post-hoc tests revealed a difference between N1 (224 Hz) and 'with' (212 Hz, $p<.001$) and between N1 and N2 (206 Hz, $p<.05$), indicating that the non-mothers produced regular, declining intonation contours with less pitch excursions than the mothers.

Forced-choice comprehension task: The ANOVA on the percent of instrument choices revealed a main effect of Addressee, $F(1,23)=11.4$, $p<.01$, indicating that there were more instrument choices in the CD sentences (57%) than in the AD sentences (52%). There was also a significant interaction between Sentence Type and Addressee, $F(1,23)=5.6$, $p<.05$, due to the fact that the difference between the two interpretations was more pronounced in the CD (instrument: 59%; modifier: 55%) than in the AD (instrument: 51%; modifier: 52%) condition. One-sample t-tests revealed that correct instrument interpretations were significantly above chance in the ambiguous and the unambiguous CD conditions (both $p's<.01$). Thus, listeners were able to identify the intended instrument interpretations reliably in CD speech suggesting that non-mothers provided clearer prosodic cues for the instrument interpretation when addressing an imaginary child.

4. GENERAL DISCUSSION

Mothers and non-mothers adopted two very different communication strategies: Non-mothers showed more pronounced prosodic disambiguation in CD speech, as indicated by noun pause durations and in comprehension data, while mothers lengthened the noun pause irrespective of targeted interpretation. In contrast, only mothers exhibited a pitch increase and the CD-typical intonation patterns thus enhancing affective prosody in CD speech while non-mothers accentuated prosodic cues to linguistic structure. We suggest that mothers favour affective prosody at the expense of linguistic prosody to attract attention and promote learning in children [13]. This is in line with evidence that children of mothers with post-natal depression, who display attenuated pitch modifications in CD speech, perform worse in associative learning tasks when maternal vocalisations serve as stimuli [5,6]. More generally, the different strategies adopted by mothers and non-mothers suggest that affective and linguistic prosody may be

independent [1,8,14,15], and that there may be a trade-off between the two systems. While the neuro-physiological evidence is still equivocal, our study supports this idea with evidence from speech production.

A potential confound in the current study is the comparison of real and imaginary CDS. Studies testing the remaining two conditions (mothers addressing imaginary children and non-mothers addressing real children) are currently underway.

REFERENCES

- [1] Baum, S., Pell, M. 1999. The neural bases of prosody: Insights from lesion studies and neuroimaging. *Aphasiology*, 13(8), 581–608.
- [2] Boersma, P., Weenink, D. 2005. *Praat: Doing Phonetics by Computer* (Version 4.3.04) [Computer Program].
- [3] Fernald, A. 1994. Human maternal vocalizations to infants as biologically relevant signals: An evolutionary perspective. In P. Bloom (Ed.), *Language Acquisition: Core Readings*. Cambridge, MA: MIT Press.
- [4] Fernald, A., Mazzie, C. 1991. Prosody and focus in speech to infants and adults. *Inf. Behav. Dev.*, 27, 209–221.
- [5] Kaplan, P., Bachorowski, J.-A., Zarlengo-Strouse, P. 1999. Child-directed speech produced by mothers with symptoms of depression fails to promote associative learning in 4-month-old infants. *Child Dev.*, 70(3), 560–570.
- [6] Kaplan, P., Bachorowski, J.-A., Smoski, M., Hudenko, W.J. 2002. Infants of depressed mothers, although competent learners, fail to learn in response to their own mothers' infant-directed speech. *Psych. Sci.*, 13(3), 268–271.
- [7] Kraljic, T., Brennan, S. 2005. Prosodic disambiguation of syntactic structure: For the speaker or for the addressee? *Cog. Psych.*, 50, 194–231.
- [8] McRoberts, G., Studdert-Kennedy, M., Shankweiler, D. 1995. The role of fundamental frequency in signalling linguistic stress and affect: Evidence for a dissociation. *Perception & Psychophysics*, (2), 159–174.
- [9] Morgan, J., Demuth, C. 1996. *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Hillsdale: Lawrence Erlbaum Associates.
- [10] Nitschke, J., Nelson, E., Rusch, B., Fox, A., Oakes, T., Davidson, R. 2004. Orbitofrontal cortex tracks positive mood in mothers viewing pictures of their newborn infants. *NeuroImage*, 21, 583–592.
- [11] Singh, L., Morgan, J., Best, C. 2002. Infants' listening preferences: Baby talk or happy talk? *Infancy* 3, 365–394.
- [12] Snedeker, J., Trueswell, J. 2003. Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *J. Mem. Lang.*, 48, 103–130.
- [13] Thiessen, E. D., Hill, E. A., Saffran, J. R. 2005. Infant-directed speech facilitates word segmentation. *Infancy*, 7(1), 53–71.
- [14] Wong, P. 2002. Hemispheric specialization of linguistic pitch patterns. *Brain Res. Bul.*, 59(2), 83–95.
- [15] Ziegler, W. 2003. Speech motor control is task-specific: Evidence from dysarthria and apraxia of speech. *Aphasiology*, 17(1), 3–36.