



Centre de Recherches
sur la Cognition
et l'Apprentissage



Centre National
de la Recherche
Scientifique



Rapport technique : 2010/02/M.AGU

Understanding Expressive Speech Acts: The Role of Prosody and Situational Context in French-Speaking 5- to 9-Year-Olds

Marc Aguert, Ludovic Le Bigot, Virginie Laval & Josie Bernicot - Centre de
Recherches sur la Cognition et l'Apprentissage (CeRCA) – CNRS - Université de
Poitiers – France

To appear in :

Aguert, M., Laval, V., Le Bigot, L. & Bernicot, J.. Understanding Expressive Speech Acts: The Role of Prosody and Context in French-Speaking 5- to 9-Year-Olds. *Journal of Speech, Language and Hearing Research*.

Address for correspondence to the first author at:

Marc Aguert,
Université de Poitiers –CNRS
Centre de Recherches sur la Cognition et l'Apprentissage (CeRCA) – CNRS UMR 6234
MSHS - Bâtiment A5 - 5 rue Théodore Lefebvre - 86000 Poitiers, France
86000 Poitiers
France
e-mail : marc.aguert@univ-poitiers.fr
Tel : 33 (0)5.49.45.48.89 ou 46.10
Fax : 33 (0)5.49.45.46.16

KEYWORDS: Understanding, Expressive Speech Acts, Prosody, Situational Context, Children.

ABSTRACT

Purpose: This study was aimed at determining the role of prosody and situational context in children's understanding of expressive utterances. Which one of these two cues will help children grasp the speaker's intention? Do children exhibit a "contextual bias" whereby they ignore prosody, like the "lexical bias" found in other studies (Friend & Bryant, 2000)?

Method: In the first experiment, a group of 5- to 9-year-old children and a group of adults performed a computerized judgment task. They had to determine the speaker's intention on the basis of an utterance produced with a particular prosody (positive or negative) in a particular situational context (positive or negative). In the second experiment, the same prosodic utterances were presented to 5- to 9-year-old children without a situational context.

Results: The 5- and 7-year-old children relied primarily on situational context, in contrast to adults who relied on prosody. The 9-year-olds relied on both cues (Experiment 1). When prosody was the sole cue (Experiment 2), all children relied on this cue to infer the speaker's intention.

Conclusions: The results are discussed and integrated into a larger conceptual framework that includes research on lexical bias and sarcasm.

INTRODUCTION

The purpose of this study was to examine the role of situational context and prosody in expressive-utterance comprehension by 5- to 9-year-old French-speaking children. A large body of research has shown that for young children, situational context plays a critical role in language comprehension, notably in cases where nonliteral language is involved (indirect requests, idioms, sarcasm, etc.). The role of prosody has been widely investigated, especially in discrepant situations, i.e., ones where lexical content conflicts with prosody. Studies on this topic have demonstrated the existence of a lexical bias: prosody carries less weight than lexical content in schoolchildren's interpretation of the speaker's intention. Although research on sarcasm has looked into the role of these two extra-lexical cues in language understanding, situational context and prosody have never been addressed from the standpoint of the strict opposition between them, independently of lexical content. Children exhibit a prosody-ignoring bias when prosody conflicts with lexical content. But what if lexical content is missing, i.e., if prosodic cues and situational cues are the only ones available? Will young children continue to show this bias in their understanding of expressive utterances? Answering this question was the main purpose of the present study. More specifically, we set out to (i) directly assess the link between situational context and prosody by controlling lexical content and thereby mirroring existing research on the contrast between lexical content and prosody, and (ii) extend the findings on this attentional bias from English to French, given that most research in this field has been conducted with English-speaking participants. This is important because it could be argued that a tendency to attend to either lexical content or situational context might be culturally-dependent.

Expressive speech acts were chosen here because they exhibit large variations in prosody, depending on the psychological state being expressed (e.g., saying "That's great" with an enthusiastic prosody or "That's sad" with a flat prosody). Expressives are speech acts whose illocutionary content is the expression of a psychological state about oneself or the world (Searle & Vanderveken, 1985). Expressives include excuses, thanks, praise, congratulations, condolences, and criticisms. Utterances like "That movie really disappointed me" or "It's great!" are expressives. Research on expressive processing by children is scarce, and the little there is concerns only expressive production, not comprehension. In addition to providing interesting material for studying comprehension due to their large intonative variations, expressives are a type of speech act that children seem to acquire at an early age (Griffiths, 1985; Halliday, 1975). Nothing is known, however, about what cues children use to understand and interpret expressive utterances.

At a more general level, the importance of situational context in children's understanding of speech is widely acknowledged today. Situational context is defined by three types of parameters: the participants's location in space and time, their characteristics, and their activities (Brown & Fraser, 1979). In particular, we know that situational context plays a fundamental role in indirect request comprehension by children as young as 1.5 years of age (Shatz, 1978) and continues to do so until at least the age of 6 or 7 (Bernicot & Legros, 1987; Spekman & Roth, 1985). The importance of situational context is not limited to requests either; it has been largely demonstrated for other kinds of language processing such as understanding promises (Astington, 1988; Laval & Bernicot, 1999) and idiomatic expressions (Laval, 2003; Levorato & Cacciari, 1992), including during adolescence, although to a lesser degree (Nippold & Martin, 1989). All of these studies oppose two types of cues, the utterance structure (linguistic forms, verb tense, linguistic conventions, etc.) and the situational context. The results invariably show that situational context has the advantage. In these studies, however, an important facet of the phonetic dimension of utterances is not considered: their prosody.

Prosody is defined as variations in the three main acoustic variables of voice: fundamental frequency, intensity, and duration (Botinis, Granström & Möbius, 2001). Prosody has several functions that facilitate utterance comprehension (Cutler, Dahan, & van Donselaar, 1997; Wells, Peppé, & Goulondris, 2004). In the present study, we focused on the pragmatic function, where prosody helps the listener understand the speaker's intention (Verschueren, 1999). Let us take an example of an interaction in a natural communication situation: *Pierre arrives at his office on Monday morning. Turning toward his colleague, he says (with a sarcastic prosody), "I had a great weekend!"* Without contextual information about the weekend, the colleague can only base her interpretation of the speaker's intention on another cue - here, prosody - to determine whether the utterance was sincere or sarcastic. In its pragmatic function, prosody is a cue that carries the illocutionary meaning of an utterance and enables the listener to determine the speaker's communicative intention (Verschueren, 1999).

To what extent are children capable of using prosody to understand an utterance? The role of prosody has often been studied in relation to that of lexical content. Lexical content includes all propositional-content information (the locutionary act; Searle, 1969), i.e., the constituent words of an utterance or "what is said". For instance, without considering situational context, the utterance "Dad gave me a new bike for my birthday" can be considered to have a positive lexical content, and "My dog ran away from home", to have a negative lexical content (Morton & Trehub, 2001). But one could just as well imagine that the new bike represents a negative event in a particular situational context, for example, if the speaker would have rather had a video game console than a new bike.

Studies on this issue began in the early seventies. It was shown using speaker's intention judgment tasks that lexical content takes precedence over prosody in schoolchildren's interpretation of utterances with contradictory

meanings (Bugental, Kaswan, & Love, 1970; Solomon & Ali, 1972). In contrast, adults base their comprehension of discrepant messages on prosody rather than lexical content. More recently, investigations into this "lexical bias" have re-emerged, with more careful experimental controls (Friend & Bryant, 2000). When judgment tasks bearing on discrepant utterances are used (e.g., happy prosody with angry content), it is consistently shown that before the age of 9 or 10, children refer to lexical content to determine the speaker's intention, even when they are explicitly instructed to pay attention to the speaker's "voice". At age 10, children start to take prosody into account, but they do not yet respond like adults (Friend, 2000; Friend & Bryant, 2000; Morton & Trehub, 2001).

This lexical bias applies to more than just those cases where affect is at stake, but covers all cases where there is a discrepancy between lexical cues and prosodic cues. For example, Moore, Harris, and Patriquin (1993) showed using certainty judgment tasks that information conveyed by prosody was secondary to information conveyed by lexical content ("think" vs. "know") for children ages 4 to 6. However, these results need to be qualified because there are different ways of observing the use of prosody at these ages. By means of a behavior-regulation task, for example, Friend (2003) found that behavior regulation by lexical content varied substantially across 4-year-olds even though, as a group, the children exhibited the characteristic bias toward lexical content. Note that in that study, prosodic cues were always associated with a facial expression that conveyed the same affective intention. When the methodology involves speech material with obscure or uninterpretable lexical content, such as utterances in a foreign language or reiterant speech (Friend & Bryant, 2000; Friend, 2000; Morton & Trehub, 2001) or low-pass filtered speech (Friend, 2000; Morton & Trehub, 2001), 4-year-olds appear to be capable of attributing an intention to the speaker based on prosody. Morton, Trehub, and Zelazo (2003) showed that following paralinguistic priming, 6-year-olds could use prosody to judge what a speaker was feeling, even when the lexical content was conflicting, whereas the children's judgments were content-based when no priming was done.

Such a prosody-ignoring bias is intriguing because prosodic cues seem to be salient and meaningful even during the earliest months of life. By 5 months of age, infants respond differently to approving and prohibiting maternal utterances (Fernald, 1993). By the time they are 7 months old, infants discriminate complex acoustic signals based on their fundamental frequency and harmonic structure (Clarkson & Clifton, 1985). Using a behavior-regulation task, Friend (2001) showed that 15-month-olds still regulated their behavior based on prosody, although a transition from affective to linguistic meaning was taking place: preverbal infants relied upon prosody to guide their behavior, but as their receptive vocabulary grew, their behavior was better regulated by lexical content. The lexical bias appeared at about 15 months.

Other studies have demonstrated the existence of a lexical bias that distinguishes adults and schoolchildren on the basis of their understanding of discrepant utterances. In this line, Eskritt and Lee (2003) showed that 3- to 5-year-olds in naturalistic conditions based their understanding of expressive utterances ("I like it", "I don't like it") on lexical content rather than on the speaker's facial expression. However, in their study, it was not stated whether the prosody was consistent with the lexical content or with the facial expression in cases where the cues were discrepant. When the inconsistency between lexical content and facial expression was made salient, more children relied on the nonlexical cue, as early as age 3. For these authors, this lexical bias is in clear contrast to the common heuristic used by adults: "Actions speak louder than words" (e.g., Blanck & Rosenthal, 1982; Demorest, Meyer, Phelps, Gardner, & Winner, 1984).

Irony and sarcasm constitute another research field in which the lexical bias has been investigated. Sarcasm is a typical case of a discrepancy between the lexical content and the speaker's intention. A speaker's intention can be ascertained by way of the situational context or the speaker's prosody, which usually do not match the utterance's

lexical content (Haiman, 1998). When they are young, children do not understand sarcasm and base their interpretations solely on lexical content (what is said, not what is meant). For adults, research has shown that prosody is the overriding cue for understanding sarcasm (Creusere, 1999), but for children, prosody does not become a critical cue until later in development. Most research has shown that situational context is the earliest cue used by children to understand sarcastic utterances. Ackerman (1982, 1983) found that 6-year-olds based their interpretation of sarcastic utterances mainly on situational context. This author also showed (Ackerman, 1983) that it was not until they were 8 years old that children became aware of prosody as a potential aid to understanding. Conversely, for Capelli, Nakagawa, and Madden (1990), prosody overrides situational context but does not appear earlier in development: prosody-based interpretations were not found until age 8, and context-based ones did not emerge until age 11. To our knowledge, the only study that has noted an early use of prosody in sarcasm understanding is Laval & Bert-Erboul (2005): using a particular type of sarcasm – sarcastic requests – these authors showed that prosody took precedence over situational context at the age of 5. This finding is of particular interest to the present study because it casts doubt on the idea that situational context necessarily takes priority in young children's speech comprehension. Laval and Bert-Erboul (2005) used a story-completion task in which two situational contexts (sarcastic or neutral) were crossed with two kinds of prosody (sarcastic or neutral). They showed that 5-year-old children understood sarcastic requests better in the sarcastic-prosody/neutral-context condition than in the neutral-prosody/sarcastic-context condition. In short, the overriding cue at age 5 was prosody. This was no longer true at age 7, where no difference between the two conditions was observed.

As we can see, the role of prosody in the understanding of utterances in discrepant situations seems to vary across studies, undoubtedly because of differences in the experimental conditions and the research issue under study. One such difference is the task used, a speaker's intention judgment task or a behavior-regulation task. Furthermore, in the research on sarcasm, the definition of sarcastic prosody (in terms of voice fundamental frequency, duration, and intensity) differs from one study to the next. One advantage of the sarcasm studies is that they operationalized situational context, something which the studies comparing only prosody and lexical content did not do, despite the fact that situation-based cues are known to be determinant for young children.

What will happen if prosody is considered alone, i.e., without being combined with lexical content? The role of prosody in language understanding has never been addressed from the standpoint of the strict opposition between prosody and situational context, independently of lexical content. We focused on a particular speech act, expressives, which are produced with large variations in prosody that depend on the psychological state being expressed.

The task used here is based on the Competition Model developed by Bates, MacWhinney, and their colleagues (Bates, MacWhinney, Caselli, Devescovi, Natale, & Venza, 1984; Bates & MacWhinney, 1987, 1989). In this model, language provides various kinds of cues – lexical, syntactic, morphological, or prosodic – to signal basic meanings, such as who is the agent of the action. The cues may be complementary, in which case they point to the same interpretation of the sentence. Cues may also be competing, pulling the interpretation in different directions, e.g., an inanimate first noun and an animate second noun. Cues also vary in terms of how consistently and reliably they relate to an interpretation (cue validity), how important that cue is to the interpretation (cue strength), and how much processing effort or time is associated with it (cue cost). Although Bates and MacWhinney's work focused on the role of animacy and word order, the argument is the same for our study, where prosody and situational context provide different cues for interpreting a sentence. The degree to which listeners rely on these cues determines their response.

Our goal in the present study was to find out which cue, prosody or situational context, is the overriding cue for 5- to 9-year-old children interpreting expressive utterances. We know that situational context is important for young children; we know that reliance on prosody is generally not observed until later in development in discrepant cases; and we know that prosody is a difficult cue to interpret for children unless it is supported by facial expressions or lexical content (Cruttenden, 1985; Bänziger & Scherer, 2005). In the light of these considerations, we hypothesized that 5- and 7-year-old children would rely mainly on situational context, unlike adults for whom prosody is the major cue. Children age 9 should fall somewhere in the middle of this developmental continuum, i.e., they should take both cues into account.

To test this hypothesis, we designed a computerized paradigm that simulates verbal interactions between two characters in nearly natural situations. The paradigm is an adaptation of the one used by Laval and Bert-Erboul (2005), who showed that prosody is the most usable cue at age 5. However, some differences between our study and theirs can explain why our hypotheses differ from Laval and Bert-Erboul's (2005). First, in our study, participants did not have to finish a story but to rate the speaker's intention on a three-point scale. Second, there was no discrepancy in the Laval and Bert-Erboul study (2005) – the situational context was neutral when the prosody was sarcastic – whereas here in our study, as in a natural interaction, the situational context was never neutral. Finally, we did not use sarcastic utterances. Prosody and lexical content were separated by using meaningless syllables instead of words to form the expressive utterances. As such, the utterances were void of lexical content and could only be interpreted on the basis of prosody, situational context, or both. This technique has been used before (Friend & Bryant, 2000), but in our study the expressive utterances were always presented in a situational context. The participants heard a conversation between two characters expressed in a "language" they did not understand. It was not a foreign language but a language of pseudo-French sentences that retained the "sound characteristics" of French. The utterances were produced with a positive or negative prosody, in a positive or negative situational context. To interpret the utterances, the participants had to determine how "positive" or "negative" the message was. The question we raised was: Do young children have a prosody-ignoring bias when prosody conflicts with situational context, as they do when it conflicts with lexical content?

EXPERIMENT 1

METHOD

PARTICIPANTS

Fifty-four native French-speaking children participated in the experiment. They were divided into three groups on the basis of age. The mean ages of the groups were 5 years 2 months (age range: 4;8–5;6 years), 7 years 0 months (*range* = 6;7-7;4), and 9 years 2 months (*range* = 8;8-9;5). Each group contained 18 children and had an equal number of girls and boys. The children were attending ordinary French public schools, and came from upper-middle and upper-class homes. They were in the normal grade for their age and had no hearing problems. A control group of 18 adult participants (9 men, 9 women) was recruited from the student body of a French university. The mean age of the adults was 23 years 9 months (*range* = 21;4-27;9). Hereafter, the four groups will be referred to respectively as the 5-year-olds, the 7-year-olds, the 9-year-olds, and the adults.

DESIGNING AND VALIDATING THE EXPERIMENTAL MATERIAL

Twelve stories about the adventures of Pilou the rabbit and Edouard the duck in everyday situations were constructed and digitized on a computer. Pictures were used to depict the communication situations, and the stories were told orally to the child by a pre-recorded off-screen voice. The stories varied according to situational context (positive vs. negative) and prosody (positive vs. negative). The situational context and prosody of the twelve stories were controlled in advance.

Controlling situational context. Sixteen situational contexts involving Pilou and Edouard in everyday situations suited to young children were generated. The situational contexts were constructed to represent a clearly positive situation (decorating a Christmas tree) or a clearly negative situation (being lost in a forest at night) for young children. In every picture, the two characters always had the same, neutral facial expression. The positiveness or negativeness of the situational contexts was pretested on ten 5-year-old children, ten 7-year-old children, and ten 9-year-old children by asking them the following question: "If you were in this situation, how would you feel, good or bad?" The twelve situational contexts that obtained a consensus at all ages considered (100% agreement) were retained for the experiment.

Controlling prosody. Again, the aim of this experiment was to study the role of situational context and prosody in children's understanding of expressive utterances, irrespective of lexical content. To this end, 40 five-syllable pseudo-utterances were generated by drawing syllables at random from a pool of randomly paired consonants and vowels. The order of the syllables in each pseudo-utterance was also random. The 40 pseudo-utterances were recorded twice in digital format by an expert adult, once with a "positive" prosody (in situations where the protagonist was feeling good) and once with a "negative" prosody (in situations where the protagonist was feeling bad). To make sure that the two types of prosody (positive and negative) were as highly contrasted as possible, the 40 pseudo-utterances were also controlled in advance by both an acoustic analysis run in PRAAT software (Boersma & Weenink, 2004) and a psychological evaluation performed by a group of adults.

Acoustic analysis. Prosody can be broadly defined in terms of three parameters: fundamental frequency (F0), intensity, and duration. However, fundamental frequency is by far the most widely studied cue in research on the verbal communication of psychological states (e.g. Lieberman & Michaels, 1962; Halliday, 1970), and it is the most important cue in the perception of prosody (Cruttenden, 1986). A low monotonous F0 pattern is characteristic of negative emotions in their "depressed" form (e.g. sadness, cold anger, disgust) whereas a high, wide-range F0 pattern is characteristic of positive emotions in their "excited" form (e.g. happiness, excitement, elation) (Scherer, 1986). The acoustic analysis on the F0 of the pseudo-utterances (PRAAT; Boersma & Weenink, 2004) showed that the positive utterances were indeed highly contrasted to the negative utterances on this parameter: the former had a high F0 mean accompanied by a large F0 range, and the latter had a lower, more monotonous F0 pattern (see Table 1).

Table 1- Acoustic Parameters of the Pseudo-Utterances

	Positive Items		Negative Items	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Duration (sec)	1.48	0.13	1.17	0.09
Maximum frequency (Hz)	372.67	15.20	165.38	26.21
Minimum frequency (Hz)	117.11	24.40	95.52	16.49
Mean frequency (Hz)	221.99	19.38	130.18	7.69
Frequency standard deviation (Hz)	72.62	4.86	10.26	3.23
Maximum intensity (dB)	79.78	0.71	77.08	0.94
Minimum intensity (dB)	45.47	2.36	45.66	5.04
Mean intensity (dB)	69.98	1.17	68.00	2.16
Intensity standard deviation (dB)	7.48	0.88	7.11	1.32

Psychological evaluation. A group of 14 naive adults (mean age: 24 years 8 months) were asked to judge the prosody of the pseudo-utterances on a 5-point Likert scale ranging from very negative (0) to very positive (5). First, the naive adults heard the set of utterances presented out of situational context. Then, on the second hearing, they had to rate the utterances. The 12 pseudo-utterances that received the most extreme ratings were retained for the experiment. The mean score of the positive utterances retained was 4.84 (*SD* = .08); it was 1.37 (*SD* = .08) for the negative utterances.



STRUCTURE OF THE EXPERIMENTAL STORIES

The twelve experimental stories varied in situational context (positive vs. negative) and prosody (positive vs. negative). Combining these two factors gave us four experimental conditions: a positive situational context with a positive prosody (C+P+), a positive situational context with a negative prosody (C+P-), a negative situational context with a negative prosody (C-P-), and a negative situational context with a positive prosody (C-P+). Two sample stories are presented in Table 2.

The stories had two parts. The first part was the situational context. The situational context was formed by a picture showing the two characters and a short oral description of the setting. For half of the stories (6) the situational context was positive (C+), and for the other half (6) it was negative (C-). The second part of the story corresponded to the production of the expressive pseudo-utterance, which was pronounced in a specific prosody. The rabbit was talking to his friend, and the first picture was shown in the background. In half of the cases the prosody of the pseudo-expressive utterance was positive (P+), and in the other half it was negative (P-). The fact that the utterances were void of lexical content was clearly explained to the children, who were simply told that Pilou lived in a faraway country and spoke a language they could not understand. Each story had three possible

endings, represented on a three-point response scale showing a picture of the rabbit crying (Pilou is feeling bad), a picture in which the rabbit's expression was hidden by his cap (We don't know how Pilou is feeling) and a picture of the rabbit with a big smile on his face (Pilou is feeling good).

Table 2 - Examples of a Story in a Positive Situational Context and a Story in a Negative Situational Context (Situational Context and Prosody are Within-Group Variables)

Story in a Positive Situational Context		
Story context		Off-screen voice: <i>Pilou and Edouard are on the merry-go-round</i>
Utterance production		Off-screen voice: <i>Pilou says to Edouard: "Déjukolumi"</i>

After the expressive utterance, the child had to answer the question "So what do you think, how is Pilou feeling?" on a 3-point scale:





Pilou is feeling good



We don't know



Pilou is feeling bad

Story in a Negative Situational Context		
Story context		Off-screen voice: <i>Pilou and Edouard are lost in the forest</i>
Utterance production		Off-screen voice: <i>Pilou says to Edouard: "Korajilamu"</i>

After the expressive utterance, the child had to answer the question "So what do you think, how is Pilou feeling?" on a 3-point scale:



Pilou is feeling good



We don't know



Pilou is feeling bad

PROCEDURE

Each child carried out the intention-judgment task individually by choosing a psychological state for the speaker in each story. The entire experimental paradigm was computerized using E-Prime software. The testing equipment included a portable computer equipped with a touch-sensitive screen, and a box with a red button on top for moving on to the next step in the story. The computerized paradigm presented the 12 stories by synchronizing the images and sounds. Each child was brought into a quiet room where, alone with the experimenter, he or she got settled in front the computer. The red-button box was placed on his or her dominant-hand side, about 20 cm from the computer screen.

Familiarization phase. During the familiarization phase, the child got accustomed to the interface and learned how to recognize the pictures corresponding to each of the three possible endings (Pilou is feeling good, Pilou is feeling bad, We don't know how Pilou is feeling). The computer randomly presented the sentences "*Pilou is feeling good*", "*Pilou is feeling bad*", and "*We don't know how Pilou is feeling*" and the child had to touch the picture on the screen that matched the psychological state expressed. When the child had learned how to do this, the experimenter ended the "game".

Experimental task. The 12 stories, displayed one by one in random order, were preceded by two practice stories that helped the child learn the task procedure. The instructions were as follows: "*Now you're going to see Pilou the little rabbit and Edouard the duck in lots of different places. You'll have to say whether Pilou is feeling good, Pilou is feeling bad, or we don't know how Pilou is feeling, by pressing on the screen as you did before. You won't be able to understand what Pilou is saying because he lives in a different, faraway country! But you're going to try to say how Pilou is feeling anyway.*" The first picture appeared when the experimenter pressed the space bar on the keyboard. Each picture was accompanied by its corresponding audio material describing the setting. Once the audio material was heard, the second picture was automatically displayed next to the first one. Then the child had to press the red button to initiate the production of the expressive utterance. At the end of the utterance, the last three pictures showing the three-point scale were displayed under the first two. At this point, the child let go of the red button and touched the screen to say whether Pilou was feeling good, feeling bad, or we don't know. Not every participant heard the same prosodic utterances with the same situational context: the utterances were presented randomly across participants. The story-presentation order was also varied randomly across participants.

CODING

Data analyses were conducted on the scores obtained by participants in each experimental condition. Recall that the participants had to choose one of three responses: "*Pilou is feeling good*", "*We don't know how Pilou is feeling*" or "*Pilou is feeling bad*". If the child thought that Pilou was feeling good or bad, he or she could choose the "*Pilou is feeling good/bad*" response. If the child didn't know how Pilou was feeling (e.g., in discrepancy cases), he or she could choose the "*We don't know how Pilou is feeling*" response. This last response was taken as an intermediate response on the three-point scale. A negative response ("*Pilou is feeling bad*") was coded 0, an indeterminate response ("*We don't know how Pilou is feeling*") was coded 1, and a positive response ("*Pilou is feeling good*") was coded 2. Responses on this scale were converted into a number and a total score was calculated for each child. This overall score thus took the full range of possible responses into account. With three stories per experimental condition, the total score fell between 0 and 6. To make the results easier to understand, we divided this score by 6 to obtain a score between 0 and 1. A score close to 1 indicated more positive responses, a score close to 0 indicated

more negative responses, and a score close to .50 indicated a lack of differentiation between positive and negative responses. Depending on the experimental condition, a score close to 0 or to 1 indicated that the participant took different cues into account. Table 3 sums up the cues used by the participant, according to the score obtained and the experimental condition.

Table 3 - Cues used by the participant depending on the score and the condition

		0	0,5	1
		←	→	→
		The participant thinks Pilou is feeling bad (using negative cues)	The participant does not know how Pilou is feeling	The participant thinks Pilou is feeling good (using positive cues)
Positive Context	Positive Prosody	-		Context, prosody, or both
	Negative Prosody		Prosody	Context
Negative Context	Positive Prosody		Context	Prosody
	Negative Prosody	Context, prosody, or both		-

The first hypothesis was that situational context is used by children at this age to understand expressive speech acts. The scores of the 5-, 7-, and 9-year-old children should be affected by the context (higher scores when the situational context is positive and lower scores when the situational context is negative). By contrast, the context should have no effect on the adults' score (scores close to .50, whatever the context). The second hypothesis was that prosody is used by 9-year-old children to understand expressive speech acts. So the scores of the 9-year-olds and the adults should be prosody-dependent (higher scores when the prosody is positive and lower scores when the prosody is negative). This should not be the case for the 5- and 7-year-olds.

RESULTS

The question studied here was what cues, context and/or prosody, are used to understand expressive utterances at the ages of 5, 7, and 9. The main data analysis was a full factorial Restricted Maximum Likelihood (RELM) analysis with an Age (4) x Prosody (2) x Context (2) mixed-model, repeated-measures design with score as the dependent variable. The two within-participant factors were context (positive context, C+; negative context, C-) and prosody (positive prosody P+; negative prosody, P-). The between-participant factor was age (5-, 7-, 9-year-old children and adults). Pairwise comparisons (Bonferroni test) were used to break down the interactions. In each experimental

condition, participants got a score between 0 and 1, which provided information about which cues they used (context, prosody, or both) to determine the speaker's intention (see Table 3).

The -2 log likelihood statistic for the model was -60.49. (ce qui indique que le modèle s'ajuste correctement aux données) Figure 1 gives the scores by age, context, and prosody.

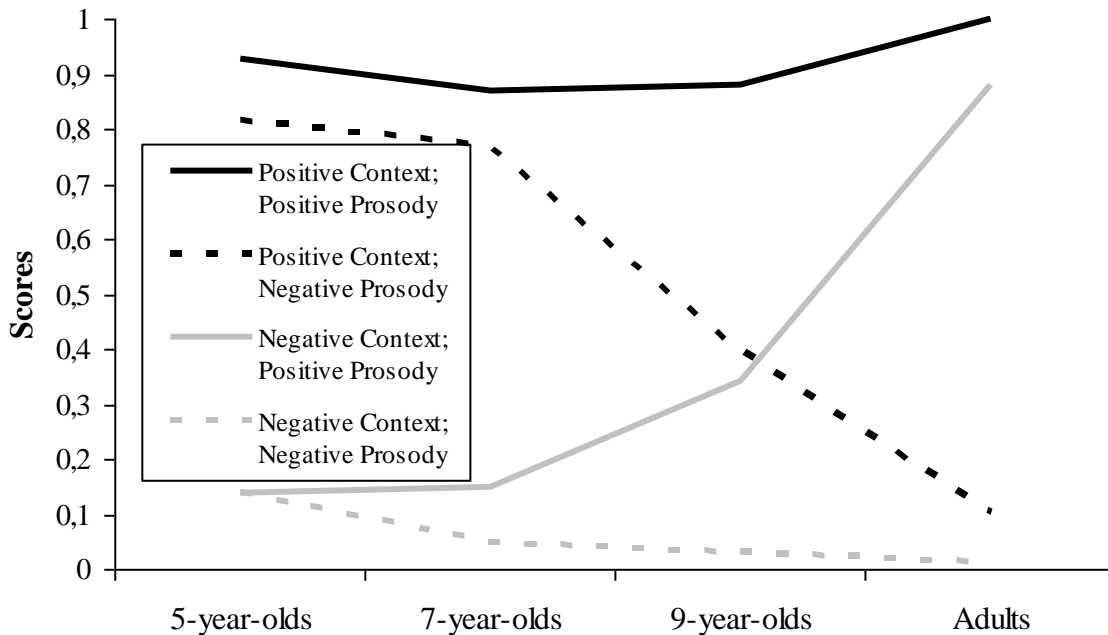


Figure 1. Score of the participants by age, context and prosody.

The analysis revealed a significant effect of context, $F(1, 191.1) = 347.58, p < .001$. The interaction between age and context was significant, $F(3, 191.1) = 29.70, p < .001$. The score was higher when the context was negative than when it was positive. Pairwise comparisons showed that 5- ($M = .87, SD = .12$), 7- ($M = .82, SD = .21$) and 9-year-olds had higher scores when the context was positive ($M_{5\text{-year-olds}} = .87, SD_{5\text{-year-olds}} = .12, M_{7\text{-year-olds}} = .82, SD_{7\text{-year-olds}} = .21, M_{9\text{-year-olds}} = .64, SD_{9\text{-year-olds}} = .22$) than when the context was negative ($M_{5\text{-year-olds}} = .14, SD_{5\text{-year-olds}} = .17, M_{7\text{-year-olds}} = .10, SD_{7\text{-year-olds}} = .13, M_{9\text{-year-olds}} = .18, SD_{9\text{-year-olds}} = .20$) ($p < .001$). There was no significant difference between the adults' scores for positive ($M_{adults} = .55, SD_{adults} = .08$) and negative ($M_{adults} = .44, SD_{adults} = .14$) contexts ($p = .63$). These results support our first hypothesis: the situational context had a significant impact on the children's score at all three ages, but not for adults.

The analysis also yielded a main effect of prosody, $F(1, 191) = 177.66, p < .001$, and an interaction between age and prosody, $F(3, 191.1) = 46.78, p < .001$. Pairwise comparisons showed that 5- and 7-year-old children did not get higher scores when the prosody was positive ($M_{5\text{-year-olds}} = .53, SD_{5\text{-year-olds}} = .13, M_{7\text{-year-olds}} = .51, SD_{7\text{-year-olds}} = .15$) than when it was negative ($M_{5\text{-year-olds}} = .48, SD_{5\text{-year-olds}} = .16, M_{7\text{-year-olds}} = .41, SD_{7\text{-year-olds}} = .18$) ($p = 1$). In contrast, 9-

year-olds and adults obtained higher scores when the prosody was positive ($M_{9\text{-year-olds}} = .61$, $SD_{9\text{-year-olds}} = .24$, $M_{adults} = .94$, $SD_{adults} = .14$) than when it was negative ($M_{9\text{-year-olds}} = .21$, $SD_{9\text{-year-olds}} = .20$, $M_{adults} = .05$, $SD_{adults} = .09$) ($p < .001$). The second hypothesis, whereby prosody has a significant impact only for 9-year-olds and adults, was supported by these results.

No other comparison reached statistical significance.

DISCUSSION

These results show that for 5- and 7-year-olds, situational context took precedence over prosody for interpreting the speaker's intention in utterances without lexical content. Conversely, adults always granted more weight to prosody. Nine-year-olds exhibited an intermediate pattern: at the group level, they relied on both cues equally. Thus, prosody emerged as a source for understanding the speaker's intention at the age of 9. This finding is consistent with most earlier studies showing an emergence of this cue around age 9, including work on the lexical bias by Solomon and Ali (1972), Friend and Bryant (2000), and Morton and Trehub (2001), and work on sarcasm by Ackerman (1983) and Capelli et al. (1990).

One of the limitations of this study was that prosodic cues were always present when the situational context was displayed (whether consistent or not). Were the 5- and 7-year-olds unaware of the prosody, or did they perceive it without making it the overriding cue? The purpose of the second experiment was to tackle this question.

EXPERIMENT 2

Based on prior research (Morton & Trehub, 2001; Friend & Bryant, 2000; Friend, 2000), the youngest children were expected to be able to infer the speaker's intention solely on the basis of prosodic cues. To test this hypothesis, we used the same experimental protocol and presented the same pseudo-utterances to the children, but without a situational context.

METHOD

PARTICIPANTS

Sixty-six native French-speaking children took part in the experiment. They were divided into three groups on the basis of age. Each group contained 22 participants and had an equal number of girls and boys (11). The three groups are hereafter called the 5-year-olds ($M_{age} = 5;0$, $range = 4;6-5;5$), the 7-year-olds ($M_{age} = 7;0$, $range = 6;8-7;5$), and the 9-year-olds ($M_{age} = 8;11$, $range = 8;5-9;5$). All of the children were from middle-class homes, were attending ordinary public schools, and were in the grade corresponding to their age.

EXPERIMENTAL MATERIAL AND PROCEDURE

The material and the procedure of Experiment 2 were the same as in Experiment 1 except that the 12 pseudo-utterances were presented without a situational context. The question the children had to answer was "How is Pilou feeling?"

RESULTS

An Age (3) x Prosody (2) repeated-measures analysis of variance (ANOVA) was conducted on the number of correct responses, i.e., judgments that accurately reflected the prosody of the utterance. Figure 2 gives the proportion of correct responses by age (5, 7, and 9) and prosody (positive and negative).

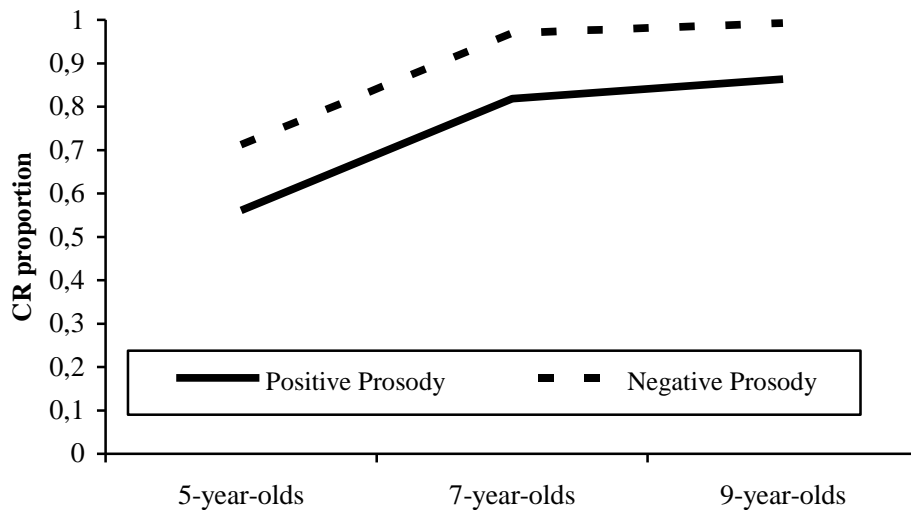


Figure 2. Proportion of correct responses (CR), by age group and Prosody (positive, negative).

The results indicated a significant effect of age, $F(2, 63) = 32.95, p < .0001, \eta^2_p = .512$. Planned comparisons showed that 7-year-old children ($M = 5.36, SD = 1.08$) gave more correct answers than 5-year-old children ($M = 3.82, SD = 1.42$), $F(1, 63) = 42.98, p < .0001$. There was no significant difference between 7-year-olds and 9-year-olds ($M = 5.57, SD = .97$), $F(1, 63) = 0.75, N.S.$ The effect of prosody was significant too, $F(1, 63) = 20.47, p < .0001$: positive prosody always triggered fewer correct responses ($M = .75, SD = 1.53$) than did negative prosody ($M = .89, SD = 1.12$). No interaction between age and prosody was observed.

Note here that the proportion of positive responses given by 5-year-olds when the prosody was positive was .56. A binomial test indicated that this score was not significantly different from chance (binomial test, $p > .19$). By

contrast, when the prosody was negative, the 5-year-olds gave significantly more positive responses than by chance (.71) (binomial test, $p < .001$).

DISCUSSION

The 7-year-olds and 9-year-olds were capable of inferring the speaker's intention on the basis of prosody. At age 5, however, prosody was interpretable only when it was negative. With similar meaningless utterances, Friend (2000) and Morton and Trehub (2001) showed that 4-year-olds were able to correctly interpret prosodic cues, for both positive and negative utterances. This earlier ability may be due to the instructions, which said to pay attention to "how the voice sounds". In the present study, prosody was not stressed by the experimenter so that the instructions would be the same as in Experiment 1. The fact that children recognized negative cues better than positive cues is consistent with several studies (Banerjee, 1997; Eskritt & Lee, 2003; Gnepp & Hess, 1986) showing that children understand situations where one must conceal negative emotions at an earlier age than situations where positive emotions are concealed, most likely due to socialization and inculcation of cultural conventions.

GENERAL DISCUSSION

The purpose of this study was to examine the role of prosody and situational context in 5- to 9-year-old children's understanding of expressive utterances, as compared to a group of adults. Again, expressive speech acts are produced very early in childhood and constitute an appropriate experimental material for studying prosody in comparison to other comprehension cues. Prosody was separated from lexical content – and this is where the specificity of our study lies – in order to clearly address the sole question of the respective roles of prosody and situational context in utterance comprehension. We designed a computerized paradigm that simulates communication situations. In Experiment 1, the two cues were always presented together; sometimes they pointed to the same interpretation of the utterance and sometimes they did not. In Experiment 2, we presented children with the same pseudo-utterances, but without a situational context, in order to measure their ability to take prosody into account when it is the only available cue.

The findings of these two experiments on the respective weights of prosody and situational context are comparable to those of previous studies comparing prosody and lexical content. When prosody is competing with another cue, young children (5- and 7-year-olds) – unlike adults who always place priority on prosody – prefer the other cue, whether it is situational context ("contextual bias") as in this study, or lexical content ("lexical bias") as in other studies (Friend & Bryant, 2000; Morton & Trehub, 2001). In the present study, when prosody was the sole available cue (Experiment 2), the children were able to use it at age 5 when it was negative and at age 7 when it had either valence, although negative prosody was more effective at helping the children grasp the speaker's intention. When a situational context was available, whether or not the children were capable of using prosody, context was the cue that led to an understanding of the speaker's intention. The fact that the children exhibited a prosody-ignoring bias is compatible with earlier findings regarding the respective weights of prosody and lexical content. Studies on this issue that have a similar methodology (a judgment task) have shown that reliance on prosody in speech comprehension emerges late in development. Children appear to base their interpretation on lexical content until they are 9 years old whenever prosody and lexical content are contradictory. In other words, between the ages of 4

and 9 years, prosody is not the overriding cue for interpreting the speaker's intention whenever it is competing with another cue, whether situational context or lexical content. The findings of the present study are consistent with the literature on this issue.

To our knowledge, among the studies on preschool and elementary school children (with the exception of studies on sarcasm), Friend's (2003) is the only one in which prosody took precedence (for some children) over another cue in discrepant situations. That study showed that some 4-year-olds used prosody to guide their behavior. However, Friend (2003) did not use an intention-interpretation task but a behavior-regulation task. This kind of task might involve different processes and may therefore lead to different results. Indeed, when judging a speaker's intention, children are mere observers of an interaction. Depending on the instructions, they must use cues that allow them to infer the mental state of another person, and this could impose constraints on their reasoning about communicative intentions (Tomasello, 1999) or on their attention flexibility (Morton & Munakata, 2002). Morton *et al.* (2003) used a rule-use paradigm to look more directly at the cognitive bases of younger children's difficulty with tasks involving conflicting information. Children heard conflicting sentences and were told first to judge them on the basis of lexical content, and then to switch and judge them on the basis of prosody. The results showed that children who noticed the conflict, and could describe it when asked, tended to switch successfully. However, unlike adults, who readily set aside salient but irrelevant cues, children seem to have difficulty altering their characteristic focus on lexical content in order to meet the demands of the task. Laval and Bert-Erboul (2005), who studied sarcastic requests, used a methodology that mixed the two approaches described above. In their study, children performed a story-completion task in which they had to finish each story by choosing one of several possible endings. The particularity of the procedure was that the endings corresponded to satisfaction or nonsatisfaction of the sarcastic request. This method is well-suited to young children because, unlike judgment tasks, it does not require an understanding of the speaker's intention, which is difficult at this stage of development. The authors found that 5-year-old children understood sarcastic requests on the basis of prosody. This early ability to use prosody might be methodology-dependent, *i.e.*, it may be due to the use of an action-based criterion to assess understanding of sarcastic requests (which is closer to a behavior-regulation task than to an intention-judgment task).

Studies using a judgment task have shown that there is a developmental transition in the use of cues to infer the speaker's intention: the overriding cue evolves from being situational context (at ages 5 and 7) to prosody (adults). This transition occurs at age 9.

The present findings for French are consistent with most previous studies showing children's tendency to ignore prosodic cues in their interpretations of the intentions of others: schoolchildren are capable of taking prosody into account, but this kind of cue is not strong enough to override lexical content or situational context. The age at which children begin to use prosodic cues is about 9 years (Solomon & Ali, 1972; Friend & Bryant, 2000; Morton & Trehub, 2001; Ackerman, 1983; Capelli *et al.*, 1990). For younger children (ages 4 and 5), prosody mainly acts as a supporting cue or modulator of lexical information (Moore, Harris, & Patriquin, 1993).

In conclusion, our results provide evidence of a developmental shift in the processing of expressive speech acts: driven at first by situational context (ages 5 and 7), expressive-interpretation processes start evolving at age 9, at which point prosody starts to carry as much weight as situational context. What we have shown (in line with Morton & Trehub, 2001; Friend, 2000) is that children ages 5 to 7 are indeed capable of interpreting prosody. However, when a situational context is available, these same children will use prosody as their first choice in determining the speaker's intention. Our study also offers further support for the recurring finding that situational

context plays a major role in speech comprehension by young children, who consistently ignore prosodic cues. This bias starts to decrease at age 9, but still exists, with a lesser impact. Research on teenagers is needed to determine the age at which prosody becomes the most reliable cue, as it is for adults. Another way to extend this study would be to use a behavior-regulation task with 4- to 10-year-old children.

REFERENCES

- Ackerman, B. P. (1982). Contextual integration and utterance interpretation: the ability of children and adults to interpret sarcastic utterances. *Child Development, 53*, 1075-1083.
- Ackerman, B.P. (1983). Form and function in children's understanding of ironic utterances. *Journal of Experimental Child Psychology, 35*, 487-508.
- Astington, J.W. (1988). Children's understanding of the speech act of promising, *Journal of Child Language, 15*, 157-173.
- Banerjee, M. (1997). Hidden emotions: preschoolers' knowledge of appearance-reality and emotion display rules. *Social Cognition, 15*, 107-132.
- Bänziger, T. & Scherer, K.R. (2005). The role of intonation in emotional expressions. *Speech Communication, 46*, 252-267.
- Bates, E., MacWhinney, B., Caselli, C., Devescovi, A., Natale, F., & Venza, V. (1984). A cross-linguistic study of the development of sentence interpretation strategies. *Child Development, 55*, 341-354.
- Bates, E., & MacWhinney, B. (1987). Competition, variation and language learning. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 157-193). Hillsdale, NJ: Erlbaum.
- Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), *The cross-linguistic study of sentence processing* (pp. 3-73). New York: Cambridge Univ. Press.
- Bernicot, J., & Legros, S. (1987). Direct and indirect directives: What do young children understand? *Journal of Experimental Child Psychology, 43*, 346-358.
- Blanck, P.D. & Rosenthal, R. (1982). Developing strategies for decoding "leaky" messages: on learning how and when to decode discrepant and consistent social communication. In R.S. Feldman (Ed.), *Development of nonverbal behaviour in children* (pp. 203-229). New York: Springer-Verlag.
- Boersma, P. & Weenink, D. (2004). *PRAAT, a system for doing phonetics by computer*. Web site: www.praat.org.
- Botinis, A., Granström, B., & Möbius, B. (2001). Developments and paradigms in intonation research. *Speech Communication, 33*, 263-296.
- Brown, P. & Fraser, C. (1979). Speech as a marker in situation. In K.R. Sherer & H. Giles (Eds.), *Social markers in speech* (pp. 33-62). Cambridge: Cambridge University Press.
- Bugental, D.E., Kaswan, J.W., & Love, L.R. (1970). Perception of contradictory meanings conveyed by verbal and nonverbal channels. *Journal of Personality and Social Psychology, 16*, 647-655.
- Capelli, C., Nakagawa, N., & Madden, C.M. (1990). How children understand sarcasm: The role of context and intonation. *Child Development, 33*, 263-296.
- Clarkson, M.G. & Clifton, R.K. (1985). Infant pitch perception: evidence for responding to pitch categories and the missing fundamental, *Journal of Acoustical Society of America, 77*, 1521-1528.
- Creusere, M.A. (1999). Theories of adults' understanding and use of irony and sarcasm: Applications to and evidence from research with children. *Developmental Review, 19*, 213-262.
- Cruttenden, A. (1985). Intonation comprehension in ten-year-olds. *Journal of Child Language, 12*, 643-61.
- Cruttenden, A. (1986). *Intonation*. Cambridge: Cambridge University Press
- Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: a literature review, *Language and Speech, 40*, 141-200.
- Demorest, A., Meyer, C., Phelps, E., Gardner, H., & Winner, E. (1984). Words speak louder than actions: understanding deliberately false remarks. *Child Development, 60*, 1527-1534.
- Eskritt, M. & Lee, K. (2003). Do actions speak louder than words? Preschool children's use of the verbal-nonverbal consistency principle during inconsistent communications. *Journal of Nonverbal Behavior, 27(1)*, 25-41
- Fernald, A. (1993). Approval and disapproval: infant responsiveness to vocal affect in familiar and unfamiliar languages, *Child Development, 64*, 657-674.
- Friend, M. (2000). Developmental changes in sensitivity to vocal paralanguage, *Developmental Science, 3(2)*, 148-162.

- Friend, M. (2001). The transition from affective to linguistic meaning. *First Language*, 21, 219-243
- Friend, M. (2003). What should I do? Behavior regulation by language and paralanguage in early childhood. *Journal of Cognition and Development*, 4(2), 161-183.
- Friend, M. & Bryant, J. (2000). A developmental lexical bias in the interpretation of discrepant messages. *Merrill-Palmer Quarterly*, 46, 140-167.
- Gnepp, J. & Hess, D.L.R (1986). Children's understanding of verbal and facial display rules. *Developmental Psychology*, 22, 103-108.
- Griffiths, P. (1985). The communicative functions of children's single-word speech. In M.D. Barrett (Ed.), *Children's single-word speech* (pp. 87-112). New York: John Wiley.
- Haiman, J. (1998). Iconicity. In J. Mey (Ed.), *Concise encyclopedia of pragmatics*. (pp. 361-366). Amsterdam: Elsevier.
- Halliday, M.A.K. (1970). *A Course in Spoken English: Intonation*. Oxford University Press: Oxford.
- Halliday, M.A.K. (1975). *Learning how to mean: Exploration in the development of language*. London: Edward Arnold.
- Laval, V. & Bernicot, J. (1999). How French-speaking children understand promises: the role of the future tense. *Journal of Psycholinguistic Research*, 28(2), 179-195.
- Laval, V. (2003). Idiom comprehension and metapragmatic knowledge. *Journal of Pragmatics*, 35, 723-739.
- Laval, V., & Bert-Erboul, A. (2005). French-speaking children's understanding of sarcasm: the role of intonation and context. *Journal of Speech, Language, and Hearing Research*, 48, 610-620.
- Levorato, M.C., & Cacciari, C. (1992). Children's comprehension and production of idioms: the role of context and familiarity. *Journal of Child Language*, 19, 415-433.
- Lieberman, P., & Michaels, S.B. (1962). Some aspects of fundamental frequency and envelope amplitude as related to the emotional content of speech. *Journal of the Acoustical Society of America*, 34, 922-927.
- Moore, C., Harris, L., & Patriquin, M. (1993). Lexical and prosodic cues in the comprehension of relative certainty. *Journal of Child Language*, 20, 153-167.
- Morton, J.B. & Munakata, Y. (2002). Are you listening? Exploring a developmental knowledge-action dissociation in a speech interpretation task. *Developmental Science*, 5, 435-440.
- Morton, J.B., & Trehub, S.E. (2001). Children's understanding of emotion in speech. *Child Development*, 72, 834-843.
- Morton, J.B., Trehub, S.E., & Zelazo, P.D. (2003). Sources of inflexibility in 6-year-olds' understanding of emotion in Speech. *Child Development*, 74, 1857-1868.
- Nippold, M.A. & Martin, S.T. (1989). Idiom interpretation in isolation versus context. A developmental study with adolescents. *Journal of Speech and Hearing Research*, 32, 58-66.
- Shatz, M. (1978). Children's comprehension of their mothers' question directives. *Journal of Child Language*, 5, 39-46.
- Spekman, N.J., & Roth, F; (1985). Children's' comprehension and production of directive forms. *Journal of Psycholinguistic Research*, 14, 331-349.
- Scherer, K.R. (1986). Vocal affect expression: a review and a model for future research, *Psychological Bulletin*, 99, 143-165.
- Searle, J.R. (1969). *Speech acts*. Cambridge: Cambridge University Press.
- Searle, J.R., & Vanderveken, D. (1985). *Foundation of illocutionary logic*. Cambridge: Cambridge University Press.
- Solomon, D. & Ali, F.A. (1972). Age trends in the perception of verbal reinforcers. *Developmental Psychology*, 7, 238-243.
- Tomasello, M. (1999). Having intentions, understanding intentions, and understanding communicative intentions. In P.D. Zelazo, J.W Astington, & D.R. Olson, (Eds.), *Developing theories of intention: social understanding and self-control* (pp. 63-75). Mahwah, NJ: Erlbaum.
- Verschueren, J. (1999). *Understanding Pragmatics*. London: Edward Arnold.
- Wells, B., Peppé, S., & Goulondris, N. (2004). Intonation development from five to thirteen. *Journal of Child Language*, 31, 749-778.

