

## **Tuning information packaging: intonational realization of topic and focus in child Dutch\***

AOJU CHEN

*Max Planck Institute for Psycholinguistics*

*(Received 26 January 2009 – Revised 2 September 2009 – Accepted 18 August 2010 –  
First published online 4 March 2011)*

### ABSTRACT

This study examined how four- to five-year-olds and seven- to eight-year-olds used intonation (accent placement and accent type) to encode topic and focus in Dutch. Naturally spoken declarative sentences with either sentence-initial topic and sentence-final focus or sentence-initial focus and sentence-final topic were elicited via a picture-matching game. Results showed that the four- to five-year-olds were adult-like in topic-marking, but were not yet fully adult-like in focus-marking, in particular, in the use of accent type in sentence-final focus (i.e. showing no preference for H\*L). Between age five and seven, the use of accent type was further developed. In contrast to the four- to five-year-olds, the seven- to eight-year-olds showed a preference for H\*L in sentence-final focus. Furthermore, they used accent type to distinguish sentence-initial focus from sentence-initial topic in addition to phonetic cues.

### INTRODUCTION

The importance of information packaging (or information structure) to communication has long been recognised (Chafe, 1976). To ensure efficient communication, the speaker forms utterances so as to structure the information he/she is attempting to convey in accordance with his/her

---

[\*] I greatly acknowledge the cooperation of the children, their parents and the teaching staff from Adalbert Primary School in Mook and De Achtsprong Primary School in Beuningen. I also would like to thank Christine Dimroth, Wolfgang Klein and Bhuvana Narasimhan for useful feedback in the course of this study; Carlos Gussenhoven for advice on intonation transcription; Marieke Hoetjes and Alice Kruisselbrink for assistance in testing the children and annotating the data; and Rik van den Brule and Steven Rekké for automating the data-processing procedure. I am also grateful to Marieke Hoetjes for serving as the second ToDI-transcriber. Preliminary results concerning the four- to five-year-old children and the adults were presented at the 16th International Congress on Phonetic Sciences in Saarbrücken, Germany, 6–10 August, 2007. Address for correspondence: Aaju Chen, Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands. e-mail: aaju.chen@mpi.nl

beliefs about the hearer's knowledge state (Clark & Haviland, 1977; Prince, 1986). There has been a long tradition to consider information packaging from a bipartite perspective. Various bipartite information packagings have been put forward in the literature, such as topic–focus, topic–comment, background–focus and theme–rheme. None of these primitive pairs alone appears to be sufficient to capture all the informational distinctions present in a sentence. To circumvent this problem, Vallduví (1992) proposed a set of tripartite primitives that were supposed to integrate all the bipartite primitives within one single schema: focus (defined as the only contribution that a sentence makes to the knowledge state of the hearer at the time of utterance), link (defined as the locus of update in the hearer's knowledge state, comparable to the notion of topic (hereafter topic)) and tail (encompassing unfocused elements that are not link-like and specify how the update should be effected). Closely related to information packaging is the information status (or cognitive status) of discourse referents (Vallduví & Engdahl, 1996). A referent can be new or given to the hearer; givenness can be of different degrees (e.g. Lambrecht, 1994). Typically, a new referent is introduced to the discourse as the focal element of a sentence but a given referent as the topic or the tail of a sentence.<sup>1</sup>

It is widely assumed that information packaging is structurally realized through syntactic operations (e.g. dislocation as a device for topic-marking in French), morphological marking (e.g. topic marker *wa* in Japanese, focus marker *ga* in Navajo) and intonation, whereas information status is marked via choice of lexical expression (e.g. an indefinite NP expressing a new referent but a pronoun expressing a given referent; see e.g. Reinhart, 1982; Prince, 1986; Gundel, 1988; Vallduví, 1992). However, strictly speaking, intonation is used in tandem with choice of lexical expression in marking the information status of discourse referents such that pronouns are spoken with little intonational prominence and full NPs can be spoken with different degrees of intonational prominence to express referents with different information statuses (Arnold, 2008). Furthermore, in the absence of lexical means, intonation is used instead for this purpose (Vallduví & Engdahl, 1996).

Children frequently produce the topic–focus structure from early on. The topic–focus structure is the most common type of successive one-word utterances in the one-word and early two-word stage, e.g. *Finger. Touch* (the child was reaching out to touch the microphone used in the recording session with her finger tip; Scollon, 1979). It is also a highly frequent type of two-word utterances in the two-word stage, e.g. *Federica acqua* 'Federica water' (The child was pretending to drink; D'Odorico & Carubbi, 2003).

---

[1] The independence of information packaging and information status can be seen in sentences where a given referent is the focus, e.g. I like HIM.

Past work has shown that children can use syntactic operations to mark topic and focus as early as 2;6 (e.g. De Cat, 2008). Furthermore, children can vary lexical expressions in an adult-like way to encode the information status of referents at the age of two or three (see Allen, Skarabela & Hughes, 2008 for an overview). These findings indicate that children under the age of four are already capable of taking their interlocutor's knowledge state into account and establishing successfully what is new or given to their interlocutor, as shown in past work on the development of theory of mind (e.g. O'Neill, 1996; De Cat, 2008).

However, surprisingly little work has been done to examine how children use intonation to realize information packaging. On the other hand, recent studies on early intonational development have shown that children acquiring Germanic and Romance languages have already developed the inventory of pitch contours in the adult model by the late two-word stage (e.g. Prieto & Vanrell, 2007; Chen & Fikkert, 2007). Further, as early as in the babbling and one-word stage children can vary the shape of pitch contours and the pitch range of the contours systematically in different emotional and interactional contexts (e.g. Flax, Lahey, Harris & Boothroyd, 1991, and references therein). Children's use of intonation to distinguish illocutionary acts (e.g. questions vs. statements) has also received a great deal of attention. Various studies in English suggest that children can use intonation to express interrogativity by the age of four but they start to use the same intonational cues as adults only at the age of seven (Patel & Grigos, 2006, and references therein). In the present study, we aim to develop our understanding of children's use of intonation to realize information packaging by investigating whether and how Dutch children use intonation to encode two information packaging primitives, topic and focus, and how they differ from adults at different developmental stages in this respect.

In the remaining part of this 'Introduction', following a brief discussion on the intonational realization of topic and focus in Dutch, we will present a review of earlier findings on children's use of intonation in topic- and focus-marking, then sketch the scope of the present study, and then propose two hypotheses on the acquisition of intonational topic- and focus-marking in Dutch children.

### *Intonational realization of topic and focus in Dutch*

It is widely accepted that intonation is the primary device for realizing information packaging in West Germanic languages. However, there is no consensus in the literature on how exactly intonation is used to mark information packaging primitives, in particular topic, in terms of either accent placement (accented or not) or accent type (the shape of the accent). More specifically, there seems to be no dispute on the claim that focus is

typically accented with a falling accent.<sup>2</sup> Topic can, however, be realized with either an accent or no accent under different conditions, according to different authors (e.g. Jackendoff, 1972; Lambrecht, 1994; Vallduví & Engdahl, 1996; Steedman, 2000). Further, the accent in the topic can be of the same type as in the focus, according to some authors (e.g. Lambrecht, 1994), but must be of a different type (e.g. a rise), according to others (e.g. Jackendoff, 1972; Steedman, 2000).

With respect to Dutch, Chen (2007) investigated the use of accent placement and choice of accent type in topic- and focus-marking in naturally spoken SVO sentences. The sentences were elicited as answers to *wh*-questions in short dialogues about either the subject or the object. She found that focus was nearly always realized with accentuation independent of its position in the sentence, and the preferred accent type was a fall. In sentence-final position, focus was also relatively frequently realized with a downstepped fall in the noun (i.e. a fall with a lower pitch peak than the highest pitch of the preceding accent). In contrast, topic was more sensitive to its position in the sentence. In sentence-initial position, topic was mostly realized with accentuation, like focus, but for rhythmic motivation in that accenting the subject noun in addition to the focused object noun led to the preferred strong–weak (in the verb)–strong rhythmic pattern (Horne, 1990). The most frequently used accent type in sentence-initial topic was also a fall. However, when a sentence-initial noun was unaccented, this occurred mostly in the topic condition. In sentence-final position, topic was preferably realized without accentuation but could be realized with a downstepped fall. Chen (unpublished observations) suggested that speakers might use a downstepped fall sentence-finally to express a lack of interest in any further discussion on the current point, following Gussenhoven, Rietveld, Kerkhoff, and Terken (2003).

*Previous studies on intonational marking of topic and focus in child language*

The few studies on children's intonational realization of information packaging primitives are mostly concerned with contrast or contrastive focus. In a picture-description task, Hornby and Hass (1970) asked English three- to four-year-olds to describe pairs of pictures that differed by one feature (subject, verb or object). They found that children frequently used contrastive stress (i.e. emphatic accentuation in the form of a fall with a rather wide pitch range) to pronounce the word carrying the contrastive information in their description of the second picture in each pair, although

---

[2] When the scope of focus is larger than one lexical word (broad focus), the accent is assigned to one of the words in focus. It is beyond the scope of the current study to discuss accent assignment in broad focus. The reader is referred to Gussenhoven (1983) and Selkirk (1995) for more discussion.

they used contrastive stress even more frequently in subject-contrast than in verb- and object-contrast. Using a similar method, MacWhinney and Bates (1978) found that the use of contrastive stress in English children was well established around the age of three but still increased between three and six. Wells, Peppé and Goulandris (2004) examined how English children used intonation to express contrast in adjective + noun phrases. They found that five-year-olds could use accentuation to mark contrast, although there was misplacement of accent in adjective-contrast. The ability to express contrast intonationally by age five has also been reported for German children. Müller, Höhle, Schmitz and Weissenborn (2006) elicited SVO sentences with a contrast either in the subject or in the object from German four- to five-year-olds by means of a question-answer task. In this task, children repeated a puppet's answer to a question about a series of comic strips; the puppet's speech lacked sentence-level prosody (i.e. intonation and rhythmical properties). Müller *et al.* found that children, like adults, uttered the words carrying contrastive information with a higher mean pitch than the words carrying no contrast with the same syntactic function and in the same sentence position (albeit differing segmentally). However, these studies tell us little about the types of accent that children use to mark contrast. Moreover, they have limited implications for children's ability to use intonation in marking non-contrastive focus.

Wieman (1976) did discuss the use of accentuation in marking non-contrastive focus. She observed in spontaneously produced two-word utterances by two-year-old English children that accent placement in the two-word stage was in the first place governed by the semantic relation between the two words. For example, in Noun/Verb-Locative utterances (e.g. *play museum*), the accent was almost always assigned to the locative (e.g. *museum*). However, the default pattern broke down if the non-default accent-bearing word carried new information or was in focus. For example, a child accented *firetruck* instead of *street* in *firetruck street* when answering his mother's question *What is in the street?* Wieman interpreted this observation as an indication that two-year-olds could strategically assign accents to mark non-contrastive focus. However, this result was based on only seven sentences in her data and its generalizability is therefore highly questionable (cf. Wells & Local, 1993). Indeed, later studies on the intonation of two-word utterances of German and Dutch two-year-olds were not able to replicate Wieman's findings. They have shown that children frequently accent both words in their two-word utterances regardless of information packaging, the syntactic organization of the utterance, and the semantic relation between the words (Behrens & Gut, 2005; Chen & Fikkert, 2007). There is thus no conclusive evidence that children can use accentuation to mark non-contrastive focus in the two-word stage. Besides, Wieman (1976) did not consider children's choice of accent type in focus-marking.

Regarding intonational topic-marking, De Ruiter (2009) examined German five- and seven-year-olds' intonational marking of preverbal topical referents in narratives elicited via a picture story-telling task. She found that children were similar to adults in the set of accent types used but different from adults in the frequencies of the accent types. Adults mostly used a steep rise with or without a preceding low-pitch stretch in the topical referents, whereas children predominantly used a weak rise. There appeared to be no systematic development between five and seven. Compared to the five-year-olds, the seven-year-olds produced more steep rises without a preceding low-pitch stretch but fewer steep rises with a preceding low-pitch stretch. However, these findings need to be verified on a larger set of data as no more than one topical NP per five-year-old and two topical NPs per seven-year-old were included in the analysis. Furthermore, as no information was available on how children would realize preverbal focal referents intonationally, it is not clear to what extent children's choice of accent type in topical referents was motivated by topic-marking or simply reflected general tendencies in production independent of information packaging.

### *The present study*

In this study, we investigated how children used accent placement and accent type to distinguish topic and focus in different positions in Dutch SVO sentences and whether there was development across time. The sentences were elicited as answers to *wh*-questions in short dialogues about either the subject referent or the object referent. Topic was operationalized as the NP expressing the referent introduced at the beginning of the dialogue and mentioned in the *wh*-question; focus was operationalized as the NP expressing the referent required via the *wh*-question. The topical referent and focal referent could be either the subject or the object of the verb, as illustrated in the short dialogues in (1). When the focal referent was the subject of the verb, the sentence had an initial focus–tail–final topic information packaging; when the focal referent was the object of the verb, the sentence had an initial topic–tail–final focus information packaging.

- (1) EXPERIMENTER: *Kijk! Een biet. Wie eet de biet?*  
 'Look! A beet. Who is eating the beet?'  
 PARTICIPANT: [*De poetsvrouw*]<sub>focus</sub> eet [*de biet*]<sub>topic</sub>.  
 'The cleaning-lady is eating the beet.'  
 EXPERIMENTER: *Kijk! Een poetsvrouw. Wat pakt de poetsvrouw?*  
 'Look! A cleaning-lady. What is the cleaning-lady picking up?'  
 PARTICIPANT: [*De poetsvrouw*]<sub>topic</sub> pakt [*een vaas*]<sub>focus</sub>.  
 'The cleaning-lady is picking up a vase.'

Children of two age groups were included in the study, four- to five-year-olds and seven- to eight-year-olds, for the following reasons. Previous studies have shown that children can use (emphatic) accentuation to mark contrastive focus by the age of four or five. It remains to be tested whether four- to five-year-olds can use intonation to encode non-contrastive focus. Furthermore, certain uses of intonation are not in place in five-year-olds but are mastered by seven- to eight-year-olds (Wells *et al.*, 2004). This implies that significant developments in aspects of intonation production have taken place between five and seven. In the context of information packaging, the question arises as to whether there are further developments in the intonational realization of non-contrastive topic- and focus-marking (hereafter topic and focus unless otherwise specified) between five and seven.

Considering that in Dutch the mapping of intonation to focus is constant, i.e. not sensitive to the position of focus in a sentence, we might assume that intonational focus-marking would be easier to acquire, compared to intonational topic-marking, which is sensitive to the position of topic in a sentence, and expect the adult-like use of intonation in focus-marking to be in place earlier than in topic-marking (Hypothesis 1). However, from a structural perspective, the intonation in the noun of the sentence-initial NP (hereafter sentence-initial noun) is less variable than the intonation in the noun of the sentence-final NP (hereafter sentence-final noun). The sentence-initial noun is mostly accented with a fall in both topic and focus, whereas the sentence-final noun is realized differently in topic than in focus. Assuming that the higher degree of variation in intonation in sentence-final nouns may pose difficulties for acquisition, we would expect the adult-like use of intonation to be acquired earlier in sentence-initial topic and focus than in sentence-final topic and focus (Hypothesis 2).

## METHOD

### *Data elicitation*

We developed a picture-matching game to elicit naturally spoken SVO sentences with either the topic–tail–focus structure or the focus–tail–topic structure in a controlled but interactive setting. In the game, the experimenter showed a child a picture each time and asked a question about the picture in order to find the matching picture. The question could be either a *who*-question or a *what*-question, as illustrated above in (1). To ensure consistency in choice of word and word order in children's answers to the questions, we introduced a few virtual robots to the game as children's informants. The robots provided children with the 'raw' answer to each question using the intended words in SVO word order via a headphone set. Crucially, the robots' sentences lacked normal sentence-level prosody

such that they were hardly imitable.<sup>3</sup> Children were supposed to reconstruct the robots' answers in their own intonation. Previous studies have shown that children's reproduction and comprehension were not affected by abnormal sentence-level prosody, although their reproduction could be affected by abnormal word order (e.g. Lahey, 1974). We therefore expected children to reconstruct the robots' sentences with no difficulty. However, to rule out misunderstandings on the task, we introduced up front the rules of the game, whereby children should use the robots' words to answer the questions but should speak like they normally did (see Appendix A for details). Further, in the light of Garrod and Pickering's (2004) alignment model for conversations, we expected children to have no difficulty with using the robots' words and word order to answer the questions. In addition, our informal survey of how children answered adult interlocutor's questions in a popular TV talk show with four- to seven-year-olds as the guests ('Praatjesmakers') revealed that children tended to answer both *wh*-questions and *yes/no* questions with full sentences or constituents larger than the words conveying the required information. This suggested that answering questions in full SVO sentences in our game should be natural for children.

Thirty-six pairs of pictures were included in the game and spread over thirty-six trials. Eighteen sentences with sentence-initial topic and sentence-final focus and eighteen sentences with sentence-initial focus and sentence-final topic were to be elicited from each participant. Each subject noun occurred in both the topic condition and the focus condition. The same held for each object noun. But each subject noun and each object noun occurred only once together in an answer sentence in the game.

### *Participants*

Twenty-eight children aged between 4;5 and 5;7 and twenty-three children aged between 7;5 and 8;10 from Adalbert Primary School in Mook and De Achtsprong Primary School in Beuningen signed up for the experiment. They were all monolingual native speakers of Dutch. Taking into account the possibility of having to exclude some data due to background noise and poor quality of segmental articulation, we decided to test all children available so that we could obtain a sufficient amount of usable data for

[3] The words that made up the robots' answer sentences were first recorded in a randomized wordlist by a female native speaker of Dutch. The words in each sentence were then spliced together with a 200 ms pause in between to form the sentence. This way, no sentence-level intonational and rhythmical properties were present in the robots' sentences. In addition, the original pitch pattern was erased and the pitch level was set at 200 Hz using Praat (Boersma, 2001) such that the sentences sounded similar in pitch level across trials. A method similar to ours was used in Müller *et al.* (2006) to create experimental materials.



reliable analysis. We aimed at getting usable data from at least ten children with no speaking or hearing deficits from each age group on the basis of recording quality, quality of articulation, and information on the children's hearing and speaking condition gathered via parent consent forms.

### *Procedure*

Children were tested individually in a quiet room at their school during school time. The experiment was conducted by two female research assistants by means of a portable set-up equipped with the NESU (Nijmegen Experiment Set-up) experiment software. One research assistant served as the experimenter each time. To help children to get comfortable with interacting with the experimenter, each session began with a short chat. Then the experimenter explained to the child how the game worked (Appendix A) and started the game with four practice trials. If the child gave elided answers or full-sentence answers but with pronouns, or appeared to try to imitate the robots' way of speaking during the practice trials, the experimenter reminded the child of the rules of the game. This intervention turned out to be necessary only in the case of a small number of the four- to five-year-olds. Most children provided full-sentence answers with no pronouns in their own intonation right from the start. No signs were observed of children having difficulty complying with the rules of the game in the experiment proper. Both experimenters were consistent in the use of prescribed intonation patterns to produce the attention getters, name the entities in the pictures, and utter the *wh*-questions throughout the experiment. They were also similar in formulation and intonation when talking about the pictures and responding to the children's remarks.

Each test session lasted about 45 minutes with the younger children and about 35 minutes with the older children. The sessions were recorded by means of a portable DAT recorder at 48 kHz sampling rate with 16-bit resolution with an external high-quality microphone placed 10–15 cm away from the children. All but one seven-year-old completed the experiment. Thirty-six responses were elicited from every child who completed the experiment. A step-by-step description of how each trial was carried out is provided in Appendix B.

## INTONATIONAL ANALYSIS

### *Data selection*

We made a first selection of the recordings on the basis of the level of background noise and quality of segmental articulation. This left us with recordings from 22 four- to five-year-olds and 18 seven- to eight-year-olds with reasonable sound quality. The recordings of 6 of the four- to

five-year-olds had to be excluded for intonation annotation for the following reasons: (i) two children turned six at the time of test; (ii) three children had minor hearing or speaking problems; and (iii) one child started to answer the experimenter's question in the middle of the robot's speech in most trials. The recordings of 7 of the seven- to eight-year-olds were excluded for intonation annotation for the following reasons: (i) three children had had minor hearing or speaking problems in the past; (ii) two children spoke Standard Dutch with a very different regional accent than the rest of the children; and (iii) two children appeared to have very unstable voice quality and this affected pitch tracking in a large number of trials. The recordings potentially suitable for intonational analysis then consisted of recordings from 16 four- to five-year-olds and 11 seven- to eight-year-olds (age range: 7;5–8;10, mean age 7;11, 6 male and 5 female).

It should be noted that 4 of the four- to five-year-olds (age range: 4;5–4;10, mean age 4;8, 1 male and 3 female) with a reasonable recording quality appeared to have a noticeably different way of speaking than the rest. Their way of speaking could be characterized with the following features: varying pitch in the range between 150 Hz and 600 Hz (vs. mostly 100–450 Hz in the other children's speech); substantial intra-speaker variations in voice quality and speaking styles so that the production from the same child sounded like coming from several children; frequently breaking down a sentence into smaller phrases; more ups and downs in their intonation; and finally sometimes laughing loudly while speaking. These four children's data were nevertheless intonationally annotated with the intention of examining whether they (hereafter the 'playful' four- to five-year-olds) differed from the four- to five-year-olds with a more consistent and everyday speaking style (hereafter the 'neutral' four- to five-year-olds) (age range: 4;5–5;6, mean age 5;2, 7 male, 5 female) in their intonational realization of information packaging. The production of the 'neutral' four- to five-year-olds was assumed to be representative of the production of average four- to five-year-olds.

### *Transcribing intonation*

*The ToDI notation.* In order to examine both the placement of accent and the type of accent involved in topic- and focus-marking, we needed to annotate both kinds of information in our data. To this end, the intonation of each sentence was transcribed following ToDI (Transcription of Dutch Intonation; see Gussenhoven, 2005; Gussenhoven *et al.*, 2003). ToDI is a notation system developed for spoken Dutch in the framework of autosegmental-metrical phonology (Ladd, 1996), the currently most widely used framework in the field of intonation phonology. It has been used to transcribe the intonation of Dutch children in the two-word stage (Chen &

Fikkert, 2007). In ToDI, the intonation contour of a sentence is described as sequences of high (H) and low (L) tones in pitch. The domain of transcription is the intonational phrase, demarcated typically by the lengthening of the phrase-final word and a pause. Generally, the pitch of an intonational phrase can start high or low and end high or low. The pitch levels at phrase boundaries are transcribed in terms of boundary tones (e.g. %L and %H at the initial boundary, L% and H% at the final boundary). Between the boundary tones, certain words bear a pitch accent (or simply an accent) and thus sound more prominent than the other words. The pitch accent is described in terms of the tones in the stressed syllable and post-stress syllable(s) of the accented word. The tone associated with the stressed syllable is demarcated with an asterisk in the labels. For example, the label H\*L (pronounced 'high star low' or 'h star l') depicts a falling contour with the fall beginning in the stressed syllable. In a disyllabic word like *poetsvrouw*, the high tone is typically realized in the stressed syllable *poets* and the low tone is realized in the post-stress syllable *vrouw*. For this reason, the high tone is also marked with an asterisk when the accent is realized in a monosyllabic word. In ToDI, five basic accent types are recognized: H\* (typically a high-pitch stretch or a weak rise), L\* (a low-pitch stretch), H\*L (a fall), L\*H (a rise), and H\*LH (a fall-rise). There are modified versions of these accents. For example, H\* and H\*L may be downstepped, and notated !H\* and !H\*L. Details on the realization of these accents in our data are given below.

Occasionally, children produced a pattern that did not fit with the description of any pitch accent type in ToDI. In this case, we gave a phonetic description of the accent using a ToDI-like label. Two labels were introduced for this purpose, H\*L^HL, and L\*HLH. The label H\*L^HL was used to describe a pattern with two distinct falls, one on each syllable, with the second fall starting noticeably higher than the first fall. The label L\*HLH was used to describe a pattern with two distinct rises, one on each syllable. Both H\*L^HL and L\*HLH occurred only in disyllabic compound nouns (e.g. *fietsclub* 'cycling-club').

*Intonation transcription.* A text-grid with two interval tiers (word tier, alternative tier), and one point tier (intonation tier) was created for each sentence and displayed together with the waveform, spectrogram and pitch contour in the TextGrid Editor window of Praat (Boersma, 2001). On the word tier, landmarks were inserted to demarcate the boundaries of each word and the words were transcribed orthographically. Word boundaries were determined mostly on the basis of the waveforms and broad-band spectrograms in combination with auditory impressions; formant contours were also used in ambiguous cases. On the intonation tier, labels of pitch accents and boundary tones were inserted following ToDI on the basis of the visual displays of the pitch contours and auditory impressions.

In ambiguous cases, informal categorization tests were done to determine the most suitable label for a certain pitch pattern. On the alternative tier, alternative labels were noted down in case of doubts.

The pitch-extraction algorithm used in Praat performs an acoustic periodicity detection on the basis of an autocorrelation method, as described in Boersma (1993). It measures  $F_0$  with an accuracy of  $10^{-6}$ , and HNR (harmonics-to-noise-ratio) values up to 60 dB and is the most accurate pitch-extraction algorithm currently available. To make sure that the pitch-extraction algorithm works in an optimal condition, it is crucial to set the base line and upper line of the pitch range in Praat adequately such that the pitch range encompasses all variations in pitch in the speaker's production. In our study, the pitch range was set between 100 Hz and 450 Hz for the 'neutral' four- to five-year-olds and the seven- to eight-year-olds, and between 150 Hz and 600 Hz for the 'playful' four- to five-year-olds. Pitch rises or falls caused by pitch tracking errors (triggered by irregularity in the voicing, creaky voice and weak voiced signals in the low pitch region) and transition from one phoneme to another (Gussenhoven, 2004: 5–10) were taken into account so that such rises and falls were not mistaken as prominence-lending pitch movements and labelled as part of an accent pattern.

*Inter-transcriber consistency.* The sentences were first transcribed by one transcriber (the author), blind to the information structural conditions. To ensure that the intonation contours were consistently and reliably transcribed, data from 6 four- to five-year-olds (from the 'neutral' speaker group) and 6 seven- to eight-year-olds were transcribed by a second transcriber, a native speaker of Dutch with extensive training in ToDI. She had no access to the first transcriber's transcription and was blind to the information structural conditions. The labels from the two transcribers were then compared and classified into three categories: (i) complete agreement; (ii) agreement in accent placement but disagreement in accent type (e.g. H\*L vs. H\*); and (iii) complete disagreement (i.e. no agreement in accent placement). Category (ii) was further divided into two sub-categories, depending on whether the labels suggested similar directions in the pitch movement (iia) (e.g. H\*L vs. !H\*L) or not (iib) (e.g. H\*L vs. H\*). The results of the comparison (Table 1) showed that the transcribers were fairly consistent.<sup>4</sup> Cases of disagreement were resolved by discussion.

*Types of accents and their realization in children's production.* The accent labels were then automatically extracted from each noun together with the orthographic transcription of the noun in each sentence by means of Praat

[4] Two of the four- to five-year-olds had a relatively narrow pitch range, which posed difficulties for transcribing the patterns in the sentence-final nouns. In a number of cases, one transcriber assigned !H\*L but the other assigned either H\*L or no accent. This caused the relatively lower percentage of complete agreement in sentence-final position in the four- to five-year-olds' data.

INTONATION OF TOPIC AND FOCUS IN CHILD DUTCH

TABLE I. *Results of the inter-transcriber comparison*

		Complete agreement	Partial agreement (direction)	Partial agreement (accent placement)	Complete disagreement	Total
4- to 5-year-olds ('neutral' group)	sentence-initial ( <i>n</i> = 155)	80.0%	2.2%	7.2%	10.5%	100%
	sentence-final ( <i>n</i> = 155)	67.5%	11.6%	12.4%	8.5%	100%
7- to 8-year-olds	sentence-initial ( <i>n</i> = 136)	92.0%	3.0%	5.0%	0	100%
	sentence-final ( <i>n</i> = 119)	92.2%	0	3.3%	4.4%	100%

scripts. The percentage distribution of each intonation pattern found in the nouns was then obtained for each speaker in SPSS.<sup>5</sup> In total, intonation patterns in 300 sentence-initial nouns and 276 sentence-final nouns produced by the 'neutral' four- to five-year-olds and intonation patterns in 59 sentence-initial nouns and 59 sentence-final nouns produced by the 'playful' four- to five-year-olds were included in the distributional count. For the seven- to eight-year-olds, intonation patterns in 277 sentence-initial nouns and 279 sentence-final nouns were included in the distributional count.

In sentence-initial nouns, six patterns occurred in the production of both the four- to five-year-olds and the seven- to eight-year-olds, i.e. H\*, H\*L, H\*L^HL, H\*LH, L\*H and 'no accent', though H\*L^HL and H\*LH were used very infrequently. In addition to these six patterns, two more accent patterns were observed occasionally in the 'playful' four- to five-year-olds' production, i.e. !H\* and L\*HLH, and one more accent pattern occurred occasionally in the seven- to eight-year-olds' production, namely, !H\*L. The H\* accent was realized as a high-pitch stretch, a weak rise, or a steep rise but with no preceding low-pitch stretch. The crucial characteristic of H\* was that the pitch either remained high across the word or reached its highest point at the end of the word. !H\* was similar to H\* but had a lower overall pitch level in comparison to the preceding high boundary tone. The H\*L accent had a falling shape, which might or might not be preceded by a rise before the fall began. The fall started either before the end of the stressed syllable or in the beginning of the post-stress syllable, and reached its lowest point within the noun, in the following unaccented verb or occasionally in the article following the verb. The !H\*L was similar to H\*L

[5] A number of nouns were not included in this count because of problems that could affect choice of intonation pattern, such as misplacement of word stress, false start, breaking a word into two parts, and phrasing (i.e. breaking a sentence into two or more intonational phrases).

but had a lower pitch peak than the preceding high boundary tone. The H\*LH accent was a fall-rise contour with an optional rise preceding the fall within the noun. The rise following the fall mostly started before the end of the noun but continued into the following unaccented verb and article. The L\*H accent was a rise contour with a distinct low-pitch stretch preceding the rise. The stylized realizations of these six accents in addition to H\*L^HL and L\*HLH are given in Table C1 (Appendix C). In Figure 1, the three most frequently used accent patterns, H\* (panel a), H\*L (panel b) and L\*H (panel c), are illustrated with realizations in the noun *paashaas* 'Easter-bunny' in the sentence *De paashaas schildert een/de voetbal* 'The Easter-bunny is painting a/the football', produced by 3 four- to five-year-olds from the 'neutral' group.

In sentence-final nouns, five patterns occurred in the production of both the four- to five-year-olds and the seven- to eight-year-olds: H\*, H\*L, !H\*L, L\*H and no accent. H\* was used occasionally. Additionally, a few instances of L\* were observed in the production of the 'neutral' four- to five-year-olds and a few instances of L\*HL and H\*L^HL were observed in the production of the 'playful' four- to five-year-olds. The H\*L accent in sentence-final nouns had a steeper falling shape than the H\*L accent in sentence-initial nouns. The !H\*L accent was realized as a fall with a pitch peak noticeably lower than the high tone in the preceding accent. The fall had already started either in the preceding word or in the stressed syllable of the sentence-final noun. The L\*HL accent was realized as a fall preceded by a low-pitch stretch in the stressed syllable. The L\* accent was realized as a low-pitch stretch. The realization of H\*, L\*H and H\*L^HL in sentence-final nouns was the same as in sentence-initial nouns. The stylized realizations of the seven accents are shown in Table C1. In Figure 2, the three most frequently used accent patterns, H\*L (panel a), !H\*L (panel b) and L\*H (panel c), are illustrated with realizations in the noun *voetbal* 'football' in the sentence *De pet schildert een/de voetbal* 'The hat is painting a/the football' produced by 3 four- to five-year-olds from the 'neutral' group.

## RESULTS AND DISCUSSION

The results are presented separately for each group of children (the 'neutral' four- to five-year-olds, the 'playful' four- to five-year-olds, and the seven- to eight-year-olds). As descriptive statistics, we report the mean percentage (%) distribution of each intonation pattern in topic and focus for each sentence position separately, because the total number of nouns differed across children in different sentence positions and the children did not use all the intonation patterns equally frequently. The mean % distribution of a given intonation pattern in a given condition was computed by averaging the percentages of the nouns spoken with the intonation pattern in the respective condition

INTONATION OF TOPIC AND FOCUS IN CHILD DUTCH

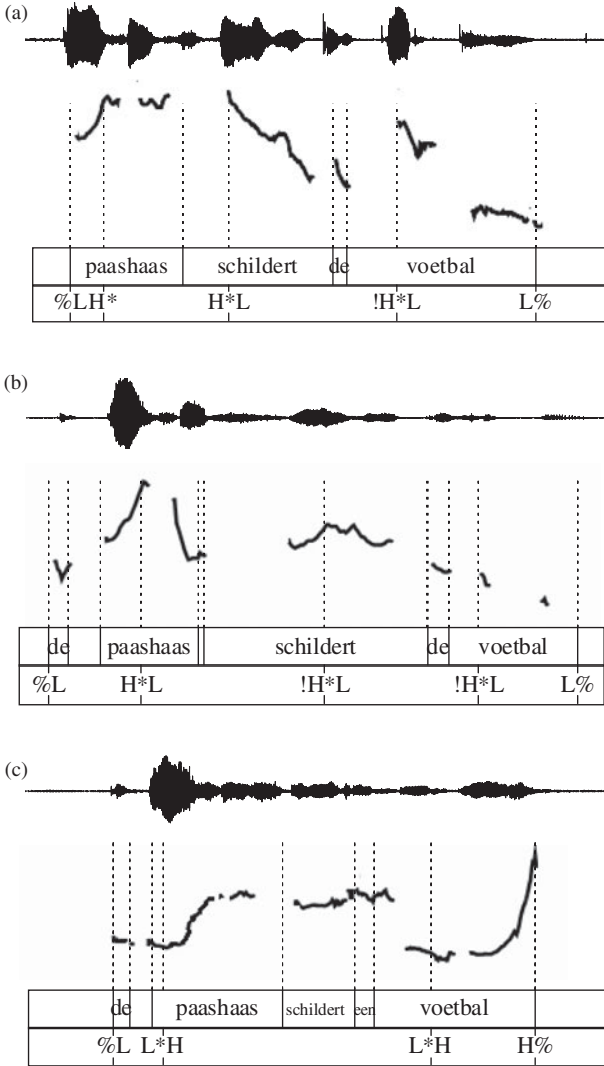


Fig. 1. Examples of H\*, H\*L and L\*H as realized in the word *paashaas* ‘Easter-bunny’ in the sentence *De paashaas schildert een/de voetbal* ‘The Easter-bunny is painting a/the football’ produced by 3 four- to five-year-old children from the ‘neutral’ group.

from all children in the respective group. For the statistical analysis, the infrequently used accent types (accent types with a mean % distribution at about 10% or lower) were merged into the category OTHER. When considering the results from the statistical analyses, we focus on the results concerning ‘no accent’ and the relatively frequently used accent types.

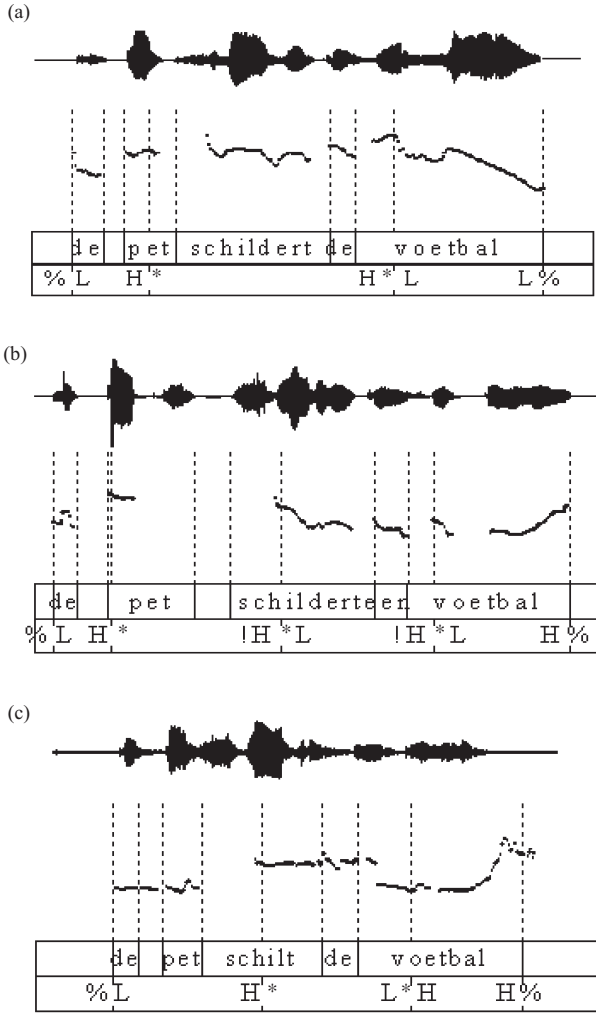


Fig. 2. Examples of H\*L, !H\*L and L\*H as realized in the word *voetbal* 'football' in the sentence *De pet schildert een/de voetbal* 'The hat is painting a/the football' produced by 3 four- to five-year-old children from the 'neutral' group. One child pronounced *schildert* as *schilt*.

*Four- to five-year-olds*

*The 'neutral' group.* As can be seen in Table 2, in sentence-initial position, both focus and topic were mostly accented, most frequently with H\*L, followed by H\*. Within the limited number of 'no accent' tokens, more occurred in the topic condition than in the focus condition. In



INTONATION OF TOPIC AND FOCUS IN CHILD DUTCH

TABLE 2. *Mean % distributions of intonation patterns in topic and focus in the 'neutral' four- to five-year-olds (N = 12)*

	H*	H*L	OTHER			no accent	Total
			H*LH	H*L^HL	L*H		
Initial focus	27.3%	55.6%	4.0%	1.3%	6.0%	5.8%	100%
Initial topic	18.9%	61.8%	2.6%	0.6%	5.1%	11.1%	100%

	H*L	!H*L	L*H	OTHER		no accent	Total
				H*	L*		
Final focus	15.5%	33.0%	40.9%	2.1%	0	8.5%	100%
Final topic	8.3%	28.5%	26.0%	2.0%	1.6%	33.6%	100%

sentence-final position, focus was mostly accented, most frequently with L\*H, followed by !H\*L and H\*L, whereas topic was most frequently realized with 'no accent'. But when topic was accented, the most frequently used accent type was L\*H, followed by !H\*L and H\*L. These observations suggested that information packaging might matter to accent placement in both sentence positions but not to choice of accent type.

To verify the observed relationship between intonation patterns and information packaging statistically, we carried out multinomial logistic regression (hereafter MLR) modelling at the significance level of 0.05 on the intonation patterns in sentence-initial nouns and sentence-final nouns separately. In each model, the independent variable (or the predictor variable) was INFORMATION PACKAGING with two categories: topic and focus. The variable SPEAKER was used to define the subgroups of the data in the model. The dependent variable (or the outcome variable) was the intonation in the nouns, consisting of four categories (H\*, H\*L, OTHER and 'no accent') in sentence-initial position and five categories (H\*L, !H\*L, L\*H, OTHER and 'no accent') in sentence-final position. The reference category was either 'no accent' or H\*L.

*Sentence-initial topic and focus.* The MLR modelling showed that the model fitting was not significantly improved after the variable INFORMATION PACKAGING was added to the model ( $-2 \log \text{likelihood} = 272.12$ ,  $\chi^2 = 6.56$ ,  $df = 3$ ,  $p = 0.087$ ). This result indicated that there was no significant overall relationship between information packaging and intonation in sentence-initial nouns. The observed tendency to use 'no accent' more frequently in topic than in focus was apparently not robust enough to establish the usefulness of INFORMATION PACKAGING in predicting choice of intonation pattern in sentence-initial position.

*Sentence-final topic and focus.* The MLR modelling showed that the model fitting was significantly improved after the variable INFORMATION PACKAGING was added to the model ( $-2 \log \text{likelihood} = 343.63$ ,  $\chi^2 = 32.97$ ,  $df = 4$ ,  $p < 0.001$ ), indicating a significant overall relationship between information packaging and intonation in sentence-final nouns. The Wald statistics (comparable to *t*-statistics in linear regression analysis) in the model with 'no accent' as the reference category showed that INFORMATION PACKAGING significantly predicted the choice between H\*L and 'no accent' ( $b = 2.06$ ,  $\text{Wald} = 17.82$ ,  $df = 1$ ,  $p < 0.001$ ), the choice between !H\*L and 'no accent' ( $b = 1.53$ ,  $\text{Wald} = 15.44$ ,  $df = 1$ ,  $p < 0.001$ ), and the choice between L\*H and 'no accent' ( $b = 1.89$ ,  $\text{Wald} = 23.1$ ,  $df = 1$ ,  $p < 0.001$ ). The odds ratios indicated that H\*L, !H\*L and L\*H were 7.83 times, 4.6 times, and 6.61 times respectively more likely to occur than 'no accent' in the focus condition compared to the topic condition. The Wald statistics in the model with H\*L as the reference category showed that there was no significant difference between H\*L and !H\*L ( $b = -0.53$ ,  $\text{Wald} = 1.65$ ,  $df = 1$ ,  $p = 0.199$ ) and between H\*L and L\*H ( $b = -0.17$ ,  $\text{Wald} = 0.17$ ,  $df = 1$ ,  $p = 0.684$ ). The odds of H\*L, !H\*L and L\*H occurring in focus was thus similar to the odds of these accent types occurring in topic. These results confirmed that accent placement differed in different conditions in sentence-final position but not choice of accent type.

*The 'playful' group.* As shown in Table 3, in sentence-initial position, focus was nearly always accented, most frequently with H\*, followed by H\*L; topic was mostly accented, most frequently with H\*L. In sentence-final position, focus was always accented, most frequently with H\*L, followed by L\*H; topic was mostly accented, most frequently with H\*L, followed by !H\*L. In both positions, tokens of 'no accent' occurred more frequently in the topic condition than in the focus condition. These observations suggested that information packaging played a role in accent placement and choice of accent type in both sentence positions.

As in the 'neutral' four- to five-year-olds' data, MLR modelling was conducted at the significance level of 0.05 on the intonation patterns in sentence-initial nouns and sentence-final nouns separately. The independent variable was INFORMATION PACKAGING with two categories, topic and focus. The variable SPEAKER was used to define the subgroups of the data in the model. The dependent variable was the intonation of the nouns, consisting of four categories (H\*, H\*L, OTHER and 'no accent') in sentence-initial position and five categories (H\*L, !H\*L, L\*H, OTHER and 'no accent') in sentence-final position. The reference category was either 'no accent' or H\*L.

*Sentence-initial topic and focus.* The MLR modelling showed that the model fitting was significantly improved after the variable INFORMATION PACKAGING was added to the model ( $-2 \log \text{likelihood} = 63.3$ ,  $\chi^2 = 14.04$ ,

INTONATION OF TOPIC AND FOCUS IN CHILD DUTCH

TABLE 3. Mean % distributions of intonation patterns in topic and focus in the 'playful' four- to five-year-olds (N=4)

	OTHER							no accent	Total
	H*	H*L	!H*	H*LH	H*L^HL	L*H	L*HLH		
Initial focus	37.1%	29.6%	2.5%	12.5%	8.3%	5.0%	0	5.0%	100%
Initial topic	8.3%	59.0%	0	9.2%	2.1%	5.8%	3.6%	11.9%	100%

	OTHER						no accent	Total
	H*L	!H*L	L*H	H*	H*L^HL	L*HL		
Final focus	44.60%	11.90%	26.50%	7.40%	1.80%	7.70%	0	100%
Final topic	31.40%	24.40%	13.30%	6.90%	0	5.00%	18.90%	100%

$df=3$ ,  $p=0.003$ ), indicating a significant overall relationship between information packaging and intonation in sentence-initial nouns. The Wald statistics in the model with 'no accent' as the reference category showed that there was no significant difference between H\*L and 'no accent' ( $b=2.2$ ,  $Wald=2.29$ ,  $df=1$ ,  $p=0.13$ ) and between H\* and 'no accent' ( $b=-1.06$ ,  $Wald=0.954$ ,  $df=1$ ,  $p=0.329$ ). The odds of H\*L occurring in focus compared to topic was similar to that of 'no accent' in that both patterns occurred more frequently in topic than in focus. Unlike H\*L and 'no accent', H\* occurred more frequently in focus than in topic on average. But this difference did not reach statistical significance. The Wald statistics in the model with H\*L as the reference category showed that INFORMATION PACKAGING significantly predicted the choice between H\* and 'H\*L' ( $b=3.25$ ,  $Wald=8.27$ ,  $df=1$ ,  $p=0.004$ ). The odds of H\* occurring in sentence-initial focus was 25.88 times higher than that of H\*L, confirming the observed preference for H\* over H\*L in sentence-initial focus.

*Sentence-final topic and focus.* The MLR modelling showed that the model fitting was significantly improved after the variable INFORMATION PACKAGING was added to the model ( $-2 \log \text{likelihood} = \chi^2 = 11.75$ ,  $df=4$ ,  $p=0.019$ ), indicating a significant overall relationship between information packaging and intonation in sentence-final nouns. The Wald statistics in the model with 'no accent' as the reference category indicated that INFORMATION PACKAGING significantly predicted the choice between L\*H and 'no accent' ( $b=3$ ,  $Wald=4.85$ ,  $df=1$ ,  $p=0.028$ ). The odds of L\*H occurring in sentence-final focus compared to sentence-final topic was twenty times higher than that of 'no accent'. The odds of !H\*L occurring in sentence-final focus relative to sentence-final topic was similar to that of 'no accent' in that both patterns occurred more frequently in topic than

TABLE 4. *Mean % distributions of intonation patterns in topic and focus in seven- to eight-year-olds (N = 11)*

	OTHER							no accent	Total
	H*	H*L	H*LH	H*L^HL	!H*L	L*H	L*HL		
Initial focus	28.6%	61.3%	1.9%	1.3%	0.7%	4.9%	0.6%	0.6%	100%
Initial topic	43.9%	43.7%	0	0.7%	0.7%	5.4%	0	5.7%	100%

	OTHER				no accent	Total
	H*L	!H*L	H*	L*H		
Final focus	40.5%	27.2%	2.8%	12.1%	17.3%	100%
Final topic	5.9%	18.8%	0	0.7%	74.5%	100%

in focus ( $b = 0.29$ ,  $Wald = 0.04$ ,  $df = 1$ ,  $p = 0.835$ ). H\*L occurred more frequently in focus than in topic on average; but this difference between H\*L and 'no accent' did not reach statistical significance ( $b = 1.95$ ,  $Wald = 2.62$ ,  $df = 1$ ,  $p = 0.106$ ). The Wald statistics in the model with H\*L as the reference category, however, showed that INFORMATION PACKAGING did not significantly predict the choice between !H\*L and H\*L ( $b = -1.6$ ,  $Wald = 3.19$ ,  $df = 1$ ,  $p = 0.074$ ) and the choice between L\*H and H\*L ( $b = -1.05$ ,  $Wald = 1.38$ ,  $df = 1$ ,  $p = 0.239$ ), suggesting inconsistency in choice of accent type between conditions among the speakers.

*Discussion.* The 'neutral' four- to five-year-olds realized focus mostly with accentuation in both sentence-initial and sentence-final position, like adults. Further, they accented sentence-initial focus mostly with H\*L and H\*, like adults. However, different from adults, they showed no preference for H\*L over the other accent types in sentence-final focus. As regards topic-marking, the 'neutral' four- to five-year-olds realized sentence-initial topic mostly with H\*L and H\* and sentence-final topic preferably with 'no accent', similar to adults. Further, when 'no accent' was used in sentence-initial position, this tended to occur more frequently in topic than in focus, as found in the production of adults.

The finding that the four- to five-year-olds used several accent types (H\*L, !H\*L, L\*H) similarly frequently in sentence-final focus cannot be explained by a failure in establishing successfully what was new or given to their interlocutor, since their use of accent placement in sentence-final position was adult-like and children of even a younger age have been shown to have a theory of mind. Rather, we suggest that the absence of a preference for H\*L in sentence-final focus was a result of several factors, i.e. children's need to express communicational purposes other than

focus-marking via a rising accent, their tendency to simply end a sentence with a rise, and the relatively frequent use of !H\*L in adult Dutch. More specifically, various non-information packaging related meanings are usually expressed in sentence-final position, such as checking and confirmation-seeking. It is possible that some of the four- to five-year-olds sometimes felt the need to check with or seek confirmation from the experimenter on the answers, especially on the answers with the new information at the end. This need overran the need to encode focus and led to the choice for L\*H instead of H\*L. Further, our informal observation suggested that some children, especially among four- to five-year-olds, habitually speak with a final rise at the end of a sentence. This accords with our impression that some of the sentences with a final rise in our data sounded like checking or confirmation-seeking but others did not. In respect of the use of !H\*L, adults use !H\*L relatively frequently in sentence-final focus, probably to express a lack of interest in further discussion on the current issue. The four- to five-year-olds might not be aware of this connotation of !H\*L, and consequently interpreted !H\*L in sentence-final focus to be functionally the same as H\*L.<sup>6</sup>

The 'playful' four- to five-year-olds realized focus predominantly with accentuation in both sentence positions, like the 'neutral' four- to five-year-olds and adults. However, they differed from the 'neutral' four- to five-year-olds and adults in accenting focus preferably with H\* in sentence-initial position. As regards topic-marking, the 'playful' four- to five-year-olds realized sentence-initial topic mostly with accentuation, like the 'neutral' four- to five-year-olds and adults. However, they also realized sentence-final topic mostly with accentuation, unlike the 'neutral' four- to five-year-olds and adults. Furthermore, even though the 'playful' four- to five-year-olds restricted the use of 'no accent' mostly to the marking of topic, they used 'no accent' much less frequently than the 'neutral' four- to five-year-olds and adults. It would thus seem that the 'playful' four- to five-year-olds preferred to accent a noun regardless of sentence position and information packaging, suggesting a tendency in the 'playful' four- to five-year-olds to speak in an emphatic manner. This tendency was also reflected in that they frequently uttered a SVO sentence as two or three intonation phrases, a typical pattern of phrasing in emphatic speech. However, the preference for H\* over H\*L in sentence-initial focus appeared to be caused

---

[6] It may be argued that the 'neutral' four- to five-year-olds did not find the game very engaging (because it was very simple) and used !H\*L to express a lack of interest in further discussion on the current point, like adults. If this were the case, we would expect the seven- to eight-year-olds to find the game even less engaging and use !H\*L more often than the 'neutral' four- to five-year-olds. As will become clear in the section on the seven- to eight-year-olds, the older children used !H\*L much less frequently than the 'neutral' four- to five-year-olds.

by a different kind of motivation because H\* is acoustically less prominent than H\*L and would sound less emphatic than H\*L. Given that the typical accent type in sentence-initial focus is H\*L, the use of H\* has an 'unconventional' ring to it. It is likely that the 'playful' four- to five-year-olds preferred to start a sentence containing new information in the initial position unconventionally.

*Seven- to eight-year-olds*

As shown in Table 4, in sentence-initial position, focus was nearly always accented, most frequently with H\*L, followed by H\*. Topic was mostly accented, most frequently with H\* and H\*L. When a sentence-initial noun was unaccented, this occurred almost only in the topic condition. In sentence-final position, focus was mostly accented, most frequently with H\*L, followed by !H\*L, whereas topic was frequently realized with 'no accent'. When topic was accented, the most frequently used accent type was !H\*L, followed by H\*L. These patterns suggested that information packaging mattered to accent placement and choice of intonation pattern in both sentence positions.

Again, MLR modelling was performed at the significance level of 0.05 on the intonation patterns in the nouns in different sentence positions separately. The independent variable was INFORMATION PACKAGING, with two categories: topic and focus. The variable SPEAKER was used to define the subgroups of the data in the model. The dependent variable was the intonation in the nouns, consisting of four categories (H\*, H\*L, OTHER and 'no accent') in sentence-initial position and four categories (H\*L, !H\*L, OTHER and 'no accent') in sentence-final position. The reference category was either 'no accent' or H\*L.

*Sentence-initial topic and focus.* The MRL modelling showed that the model fitting was significantly improved after the variable INFORMATION PACKAGING was added to the model ( $-2 \log \text{likelihood} = 219.33$ ,  $\chi^2 = 12.66$ ,  $df = 3$ ,  $p = 0.005$ ), indicating a significant overall relationship between information packaging and intonation in sentence-final nouns. The Wald statistics in the model with 'no accent' as the reference category indicated that INFORMATION PACKAGING significantly predicted the choice between H\*L and 'no accent' ( $b = 2.36$ ,  $\text{Wald} = 4.77$ ,  $df = 1$ ,  $p = 0.029$ ). The odds of H\*L occurring in sentence-initial focus compared to sentence-initial topic was 10.62 times higher than that of 'no accent'. The odds of H\* occurring in sentence-initial focus compared to sentence-initial topic was similar to that of 'no accent' in that both patterns occurred more frequently in topic than in focus ( $b = 1.70$ ,  $\text{Wald} = 2.42$ ,  $df = 1$ ,  $p = 0.12$ ). The Wald statistics in the model with H\*L as the reference category showed that INFORMATION PACKAGING significantly predicted the choice between H\* and H\*L

( $b = -0.67$ , Wald = 6.37,  $df = 1$ ,  $p = 0.012$ ). The odds of H\*L occurring in sentence-initial focus compared to sentence-initial topic was 1.95 times higher than that of H\*, suggesting a preference for H\*L in marking sentence-initial focus.

*Sentence-final topic and focus.* The MLR modelling showed that the model fitting was significantly improved after the variable INFORMATION PACKAGING was added to the model ( $-2 \log \text{likelihood} = 286.28$ ,  $\chi^2 = 115.7$ ,  $df = 3$ ,  $p < 0.001$ ), indicating a significant overall relationship between information packaging and intonation in sentence-final nouns. The Wald statistics in the model with 'no accent' as the reference category showed that INFORMATION PACKAGING significantly predicted the choice between !H\*L and 'no accent' ( $b = 1.93$ , Wald = 31.68,  $df = 1$ ,  $p < 0.001$ ) and between H\*L and 'no accent' ( $b = 3.32$ , Wald = 64.03,  $df = 1$ ,  $p < 0.001$ ). The odds of !H\*L and H\*L occurring in sentence-final focus relative to sentence-final topic was 6.89 times and 27.68 times respectively higher than that of 'no accent', suggesting a preference for accenting focus in sentence-final position. The Wald statistics in the model with H\*L as the reference category showed that INFORMATION PACKAGING significantly predicted the choice between !H\*L and H\*L ( $b = -1.39$ , Wald = 10.74,  $df = 1$ ,  $p < 0.001$ ). The odds of H\*L occurring in sentence-final focus relative to sentence-final topic was four times higher than that of !H\*L, suggesting a preference for accenting sentence-final focus with H\*L.

*Discussion.* Like adults, the seven- to eight-year-olds realized focus and sentence-initial topic predominantly with accentuation, and sentence-final topic mostly without accentuation. Further, the seven- to eight-year-olds showed an adult-like preference for H\*L over !H\*L in marking sentence-final focus. This was a significant development compared to the four- to five-year-olds, who showed no preference for a certain accent type. Interestingly, the seven- to eight-year-olds preferred H\*L to H\* in sentence-initial focus but not in sentence-initial topic, unlike adults, who favoured H\*L in both cases. In other words, the seven- to eight-year-olds used accent type to distinguish focus from topic in sentence-initial position. At first glance, this result looked striking, given that the four- and five-year-olds were adult-like in predominantly using H\*L in both topic and focus in sentence-initial position. However, considering that adults varied the duration, pitch minimum and pitch span of H\*L to distinguish topic and focus in sentence-initial position but the four- to five-year-olds did not do so (Chen, 2009), the use of accent type for this purpose could be considered another development in the seven- to eight-year-olds. As the seven- to eight-year-olds only varied the pitch minimum and pitch span of H\*L to distinguish sentence-initial focus from sentence-initial topic (Chen, 2009), choice of accent type could serve as an additional strategy.

## GENERAL DISCUSSION

Our results show that the four- to five-year-olds were adult-like in intonational topic-marking. Regarding focus-marking, they were adult-like in the use of accent placement in both sentence positions and choice of accent type in sentence-initial position but not in choice of accent type in sentence-final position (i.e. showing no preference for H\*L). These results lend support to Hypothesis 2, which predicted earlier mastery of intonational topic- and focus-marking in sentence-initial position than in sentence-final position. The non-adult-like use of accent type in sentence-final focus could be attributed to several factors, including the need for seeking confirmation and checking, the tendency to end a sentence with a final rise – both leading to the frequent use of L\*H – and the relatively frequent use of !H\*L in sentence-final focus in adult Dutch. The use of accent type was further developed between five and seven. In contrast to the four- to five-year-old children, the seven- to eight-year-old children showed a preference for H\*L in sentence-final focus. Furthermore, they used accent type to distinguish sentence-initial focus from sentence-initial topic in addition to varying the pitch minimum and pitch span of H\*L.

In the comparison between the ‘neutral’ four- to five-year-olds and the ‘playful’ four- to five-year-olds, we have observed systematic differences in how the two groups of four- to five-year-olds realized topical and focal nouns intonationally. The ‘playful’ four- to five-year-olds’ choice of intonation pattern appeared to be driven by the intention to sound emphatic and unconventional, instead of the form–function mappings between intonation and information packaging in adult Dutch. Individual variation in children’s intonational skills at a given developmental point has been reported for both production and comprehension tasks in earlier work (e.g. Wells *et al.*, 2004). However, nothing is known about individual variation in the rate and route of acquisition longitudinally. In the case of the ‘playful’ four- to five-year-olds in our study, it is possible that they will grow out of the ‘playful’ stage when they get a bit older, and acquire the adult-like use of intonation in topic- and focus-marking later than the ‘neutral’ four- to five-year-olds. It is also possible that the ‘playful’ four- to five-year-olds may undergo a different developmental path to the adult-like use of intonation than the ‘neutral’ four- to five-year-olds. Future longitudinal work is needed to find out how such differences among children affect the rate and route of acquiring intonational topic- and focus-marking.

Our findings also have useful implications for research on argument realization in child language. As mentioned in the ‘Introduction’, studies in this area have focused on the relationship between the information status of a referent and choice of lexical form to express the referent. The general finding is that children use relatively more informative lexical forms (e.g. full NPs) to express referents with a low degree of givenness



and less informative lexical forms (e.g. pronouns) to express referents with a high degree of givenness from early on. When using full NPs to express new referents, children under the age of five predominantly use the definite NPs, going against the general pattern of indefiniteness–newness and definiteness–givenness in adult speech (e.g. Maratsos, 1974). However, no known study has investigated whether children vary intonation in definite NPs to distinguish differences in information status. There has also been no research done on children's use of intonation when a less informative lexical form is used in a referentially ambiguous context. A case in point is the use of pronouns to express referents that have potential competitor referents, as in *The camel hit the lion and then he (the lion) hit the elephant* (Solan, 1980) and *Goofy is petting Bart. Now you pet him (Goofy)* (McDaniel & Maxfield, 1992). In such cases, the pronouns *he* and *him* should be spoken with emphatic accentuation. Studies on children's comprehension of pronouns have shown that even five-year-olds fail to interpret the accented pronouns correctly. This result has been considered in the literature as evidence showing that children's comprehension of the accentuation lags behind their use of accentuation in production (e.g. Cutler & Swinney, 1987). However, it remains to be investigated whether children can use accentuation in pronominal disambiguation in production in the first place.

A final issue for further research is concerned with language-specificity in the intonational marking of information packaging. In languages like Dutch, choice of accent type plays an important role in encoding focus. In tone languages like Mandarin Chinese, the shape of the pitch contour is lexically determined and the intonational encoding of topic and focus is mainly implemented as variations in the maximum pitch and the minimum pitch of the topical word, the focused word and the immediately following word, which is neither topical nor focused (Wang & Xu, 2006). The question arises as to how Mandarin Chinese-speaking children differ from Dutch-speaking children in the developmental path to adult-like intonational realization of information packaging. Future work can be directed to comparisons between children from typologically different languages to gain an insight into language-specific challenges in the acquisition of the intonational realization of information packaging in spoken discourse.

## APPENDIX A

The picture-matching game was described to each child as follows:

I've got two boxes full of pictures here. The pictures are about a fantasy world where anything goes. For example, animals can speak like humans and tables can kick objects with their legs. A picture from one box goes

together with a picture from the other box. But the pictures have got mixed up and I need your help to sort the pictures out. Every time, I take a picture from one box and ask you a question about it. Of course you don't know the answer either. But you will get help from a robot. The robot will tell you the answer via a headphone set so that only you can hear what the robot says. You will then know the answer and be able to answer my question. There will be four robots to help you in the game. The robots always know the answers but they speak in a very funny way. You will use the robots' words to answer my questions but you should speak like you normally do. These are the rules of the game. With the help of your answers, I can then find the matching pictures.

## APPENDIX B

In the picture-matching game, each trial was carried out in the following steps. First, the experimenter took a picture (e.g. a picture of a cleaning-lady) from one box, drew the child's attention to the picture, and established what the picture was by saying *Kijk! Een poetsvrouw* 'Look! A cleaning-lady!' with a high-pitch stretch or a weak rise on the attention getter *Kijk*, and a rise on the noun to indicate that more was to come following the naming. These intonation patterns are common in comparable conversational situations in Dutch. Often the child chatted a bit about the picture with the experimenter. If the child did not do so but the word was identified as a potentially difficult word for four- to five-year-olds in our pilot tests, the experimenter took the initiative to talk a bit about the picture. This was done to make sure that the entity on the picture was referentially given to the child by the time a question about the entity was asked.

The experimenter then asked a question about the already introduced referent (e.g. In the picture, the cleaning-lady seemed to be picking up something; the question was then *Wat pakt de poetsvrouw?* 'What is the cleaning-lady picking up?'), with a high-pitch stretch or a weak rise on the *wh*-word, a fall or a downstepped fall on the verb, an optional downstepped fall on the noun, and a low tone at the end of the *wh*-question. An informal survey of *wh*-questions uttered as information-seeking questions in the Spoken Dutch Corpus (Oostdijk, 2000) showed that these patterns are commonly used in *wh*-questions in spoken Dutch.

Following the experimenter's question, the child turned to a robot for help by clicking on a picture of the robot displayed on his or her computer screen. The child received the answer from the robot via a headphone set. Four robots were used in the game; they were randomly assigned to the trials. Having heard what the robot said, the child answered the experimenter's

question using the same words as the robot but in his or her own intonation (e.g. *De poetsvrouw pakt een vaas*. ‘The cleaning-lady is picking up a vase.’).

Finally, the experimenter looked for the matching picture (e.g. the picture of a vase) in the other box and handed over both pictures to the child.

APPENDIX C

TABLE C1. Stylized realizations of accent types in disyllabic words in children’s production. The filled square stands for the stressed syllable; the empty square stands for the following unstressed syllable. The height of the square represents the pitch range within which an accent is realized. The dotted lines represent additional realizations of an accent ending at different pitch heights at the end of a phrase. The non-standard ToDI accent types are in italics in the leftmost panel.

	in sentence-initial nouns			in sentence-final nouns	
<b>H*</b>					
<b>!H*</b>					
<b>H*L</b>					
<b>!H*L</b>					
<b>H*LH</b>					
<b>L*</b>					
<b>L*H</b>					
<b>L*HL</b>					
<i>H*L^HL</i>					
<i>L*HLH</i>					

## REFERENCES

- Allen, S. E. M., Skarabela, B. & Hughes, M. (2008). Using corpora to examine discourse effects in syntax. In H. Behrens (ed.), *Corpora in language acquisition research: Finding structure in data*. Amsterdam: Benjamins.
- Arnold, J. E. (2008). Reference production: Production-internal and addressee-oriented processes. *Language and Cognitive Processes* 23(4), 495–527.
- Behrens, H. & Gut, U. (2005). The relationship between prosodic and syntactic organization in early multiword speech. *Journal of Child Language* 32, 1–34.
- Boersma, P. (1993). Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. In *Proceedings of the Institute of Phonetic Sciences*, Amsterdam.
- Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glott International* 5(9/10), 341–45.
- Chafe, W. C. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. Li (ed.), *Subject and Topic*. New York: Academic Press.
- Chen, A. (2007). Intonational realisation of topic and focus by Dutch-acquiring 4- to 5-year-olds. In J. Trouvain & J. B. William (eds), *Proceedings of the 16th International Congress of Phonetic Sciences*. Pirrot GmbH: Dudweiler.
- Chen A. (2009). The phonetics of sentence-initial topic and focus in adult and child Dutch. In M. Vigário, S. Frota & M. J. Freitas (eds), *Phonetics and phonology: Interactions and interrelations*. Amsterdam: Benjamins.
- Chen, A. & Fikkert P. (2007). Intonation of early two-word utterances in Dutch. In J. Trouvain & J. B. William (eds), *Proceedings of the 16th International Congress of Phonetic Sciences*. Pirrot GmbH: Dudweiler.
- Clark, H. H. & Haviland, S. E. (1977). Comprehension and the given-new contact. In R. O. Freedle (ed.), *Discourse production and comprehension*. Hillsdale, NJ: Erlbaum.
- Cutler, A. & Swinney, D. A. (1987). Prosody and the development of comprehension. *Journal of Child Language* 14, 145–67.
- De Cat, C. (2008). Experimental evidence for preschoolers' mastery of 'topic'. In A. Gavarró & M. J. Freitas (eds), *Proceedings of GALA 2007*. Cambridge: Cambridge Scholars Publishing.
- De Ruiter, L. E. (2009). The prosodic marking of topical referents in the German 'Vorfeld' by children and adults. *The Linguistic Review* 26(2/3), 329–54.
- D'Odorico, L. & Carubbi, S. (2003). Prosodic characteristics of early multi-word utterances in Italian children. *First Language* 23(1), 97–116.
- Flax, J., Lahey, M., Harris, K. & Boothroyd, A. (1991). Relations between prosodic variables and communicative functions. *Journal of Child Language* 18, 3–19.
- Garrod S. & Pickering M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences* 8, 8–11.
- Gundel, J. K. (1988). Universals of topic-comment structure. In M. Hammond, E. Moravcsik & J. R. Wirth (eds), *Studies in syntactic typology*. Amsterdam: Benjamins.
- Gussenhoven, C. (1983). Testing the reality of focus domains. *Language and Speech* 26, 61–80.
- Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
- Gussenhoven, C. (2005). Transcription of Dutch intonation. In S. Jun (ed.), *Prosodic typology and transcription: A unified approach*. Oxford: Oxford University Press.
- Gussenhoven, C., Rietveld, R., Kerkhoff, J. & Terken, J. (2003). Transcription of Dutch intonation: Courseware, 2nd edn, at <http://todi.let.kun.nl/ToDI/home.htm>.
- Horne, M. (1990). Accentual patterning in 'new' vs. 'given' subjects in English. *Lund Working Papers in Linguistics* 36, 81–97.
- Hornby, P. A. & Hass, W. A. (1970). Use of contrastive stress by preschool children. *Journal of Speech and Hearing Research* 13, 359–99.
- Jackendoff, R. (1972). *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

- Ladd, D. R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.
- Lahey, M. (1974). Use of prosody and syntactic markers in children's comprehension of spoken sentences. *Journal of Speech and Hearing Research* **17**, 656–68.
- Lambrecht, K. (1994). *Information structure and sentence form: Topics, focus, and the representations of discourse referents*. Cambridge: Cambridge University Press.
- MacWhinney, B. & Bates, E. (1978). Sentential devices for conveying givenness and newness: A cross-cultural developmental study. *Journal of Verbal Learning and Verbal Behavior* **17**, 539–58.
- Maratsos, M. (1974). Preschool children's use of definite and indefinite articles. *Child Development* **45**, 446–55.
- McDaniel, D. & Maxfield, T. L. (1992). Principle B and contrastive stress. *Language Acquisition* **2**(4), 337–58.
- Müller, A., Höhle, B., Schmitz, M. & Weissenborn, J. (2006). Focus-to-stress alignment in 4- to 5-year-old German-learning children. In A. Belletti, E. Bennati, C. Chesi, E. Di Domenico & I. Ferrari (eds), *Proceedings of GALA 2005*. Cambridge: Cambridge Scholars Press.
- O'Neill, D. K. (1996). Two-year-old children's sensitivity to a parent's knowledge stage when making requests. *Child Development* **67**(2), 659–88.
- Oostdijk, N. H. J. (2000). The Spoken Dutch Corpus: Outline and first evaluation. In *Proceedings of the 2nd International Conference on Language Resources and Evaluation*.
- Patel, R. & Grigos M. I. (2006). Acoustic characterization of the question–statement contrast in 4-, 7- and 11-year-old children. *Speech Communication* **48**, 1308–318.
- Prieto, P. & Vanrell, M. M. (2007). Early intonational development in Catalan. In J. Trouvain & J. B. William (eds), *Proceedings of the 16th International Congress of Phonetic Sciences*. Pirrot GmbH: Dudweiler.
- Prince, E. F. (1986). On the syntactic marking of presupposed open propositions. In A. Farley, P. Farley & K. E. McCullough (eds), *Papers from the Parasession on Pragmatics and Grammatical Theory*, 22nd Regional Meeting, Chicago Linguistic Society.
- Reinhart, T. (1982). Pragmatics and linguistics: An analysis of sentence topics. *Philosophica* **27**, 53–94.
- Scollon, R. (1979). A real early stage: An unzipped condensation of a dissertation on child language. In E. Ochs & E. Schieffelin (eds), *Developmental pragmatics*. New York: Academic Press.
- Selkirk, E. O. (1995). Sentence prosody: Intonation, stress and phrasing. In J. Goldsmith (ed.), *Handbook of phonological theory*. London: Blackwell.
- Solan, L. (1980). Contrastive stress and children's interpretation of pronouns. *Journal of Speech and Hearing Research* **23**, 688–98.
- Steedman, M. (2000). Information structure and the syntax–phonology interface. *Linguistic Inquiry* **31**(4), 649–89.
- Vallduví, E. (1992). *The informational component*. New York: Garland.
- Vallduví, E. & Engdahl, E. (1996). The linguistic realisation of information packaging. *Linguistics* **34**, 459–519.
- Wang, B. & Xu, Y. (2006). Prosodic encoding of Topic and Focus in Mandarin. In R. Hoffmann & H. Mixdorff (eds), *Proceedings of Speech Prosody 2006*. Dresden: TUD Press.
- Wells, B. & Local, J. (1993). The sense of an ending: A case of prosodic delay. *Clinical Linguistics and Phonetics* **7**(1), 59–73.
- Wells, B., Peppé, S. & Goulandris, N. (2004). Intonation development from five to thirteen. *Journal of Child Language* **31**, 749–78.
- Wieman, L. A. (1976). Stress patterns in early child language. *Journal of Child Language* **3**, 283–86.