Acoustic emphasis in four year olds

Elizabeth Wonnacott and
Department of Experimental Psychology, University of Oxford

Duane G. Watson
Department of Psychology, University of Illinois – Urbana Champaign

Abstract

Acoustic emphasis may convey a range of subtle discourse distinctions, yet little is known about how this complex ability develops in children. This paper presents a first investigation of the factors which influence the production of acoustic prominence in young children's spontaneous speech. In a production experiment, SVO sentences were elicited from 4 year olds who were asked to describe events in a video. Children were found to place more acoustic prominence both on 'new' words and on words which were 'given' but had shifted to a more accessible position within the discourse. This effect of accessibility concurs with recent studies of adult speech (Dahan, Tanenhaus, & Chambers, 2002, Watson, Arnold & Tanenhaus 2005). We conclude that, by age four, children show appropriate, adult-like use of acoustic prominence, suggesting sensitivity to a variety of discourse distinctions.

Keywords

Language Development; Language Production; Prosody; Accent; Discourse; Accessibility

---

In adult speech, certain words may be emphasized over others, using some combination of increased pitch, higher intensity and longer duration in order to signal information about discourse structure. However, little is known about how this ability develops in children. In this paper, we investigate the relationship between acoustic prominence and discourse structure in four-year olds' spontaneous production.

For adult speech, a standard hypothesis has been that acoustic prominence[1] is a function of the *givenness* of information in the discourse: information that is new is emphasized while information that is given is not (Brown, 1983;Chafe, 1974;Prince, 1981). However, more recently, researchers have argued that the use of emphasis may reflect more fine-grained discourse distinctions. One notion that has played a central role in understanding discourse status is that of the *accessibility* of information within the discourse, i.e. the degree to which information is activated or salient. Information that has been mentioned frequently, or that is highly topical, is more accessible than information that is new, or information that has been mentioned previously but only peripherally. This property has been shown to affect word form

---

Corresponding Author: Elizabeth Wonnacott, Department of Experimental Psychology, University of Oxford, Oxford, OX1 1UD, Email: elizabeth.wonnacott@psy.ox.ac.uk, Phone: 01865 271319.

[1]We are focusing on what have traditionally been called 'presentational accents' (or H* in ToBI notation). However accenting is often discussed as a categorical phenomenon, and some researchers have argued that distinctions in the salience of a word may be continuous (e.g. see Ladd, 1996 for a discussion). We therefore use the more neutral term 'acoustic prominence' when discussing differences in salience.

choice and pronoun usage: referents that are highly accessible are typically referred to using pronouns and less accessible referents are typically referred to using full noun phrases (Ariel, 1990;Grosz & Sidner, 1986;Gundel, Hedberg, & Zacharski, 1993).

Dahan, Tanenhaus and Chambers (2002) proposed that the relationship between word form choice and accessibility extends to the use of acoustic prominence such that information that shifts from a low to a high level of accessibility is more likely to be produced with an accent than information that remains highly accessible in a discourse. To test this hypothesis, they examined how listeners interpreted accented information. In an eye-tracking study, participants were given instructions to move objects to new locations on a computer display. In line with the notion of a 'given-new' distinction, they found that when the object to be moved was accented, participants were faster at fixating referents that had not been mentioned than those that had been mentioned and had moved previously. However, participants were also faster at fixating an accented referent if the object had been mentioned but had not moved (e.g. it had been referred to as a landmark and so was not highly accessible). These findings suggest that listeners interpret acoustic prominence as signals about accessibility change.

Watson, Arnold and Tanenhaus (2005) found similar patterns in language production. Speakers described the movement of objects on a computer screen to a partner who had to mimic the actions, producing sentences like the ones below in (1):

(1) a. *Given-NonShift* Condition, e.g.

Put the **bed** above the house.

Put the **bed** above the pineapple.

b. *Given-Shift* Condition, e.g.

Put the house above the **bed**.

Put the **bed** above the pineapple.

c. *New-Condition*, e.g.

Put the house above the bell.

Put the **bed** above the pineapple.

The critical word in each trial was the direct object of the verb in the second sentence. Speakers used the most acoustic prominence when referring to referents that were new, (as in 1c), and given referents that had shifted accessibility in the discourse (1b). Speakers used less prominence on the critical word when the referent had not shifted its discourse status (1a).

With a few exceptions (see Snow & Balog, 2002), little is currently known about how the ability to use acoustic prominence develops in children, or how children might differ from adults. The question is important since sensitivity to discourse structure may not be fully developed in young children (e.g. Avrutin, 1999; Arnold, Novick, Brown-Schmidt, Eisenband & Trueswell, 2001). However, recent work on preschoolers' pronoun use suggests at least some sensitivity to discourse accessibility in both comprehension and production. Referents that have occupied subject position previously in a discourse, a position which is associated with high accessibility, are more likely to be referred to with a pronoun by 4-year olds (Hickmann & Hendriks, 1999). Children at ages 2.5 to 3.5 are more likely to respond to a question with an answer that includes a pronoun if the referent was mentioned in the question (Campbell, Brooks, & Tomasello, 2000). 3-year old children are more likely to interpret a pronoun as referring to the subject than to the object (Song & Fisher, 2005). In sum, these findings suggest an ability to track shifts in discourse accessibility. We now ask whether these same discourse distinctions are reflected in their use of acoustic prominence.

In this experiment, we attempted to elicit pairs of SVO sentences from descriptions of events in a video. The agent of the second sentence was the target word and the first sentence established a discourse context. There were three trial types:

(2) a. Given–NonShift, e.g.

Sentence 1: The **lion** hit the giraffe.

Sentence 2: The **lion** hit the elephant.

b. Given–Shift, e.g.

Sentence 1: The giraffe hit the **lion**.

Sentence 2: The **lion** hit the elephant.

c. New, e.g.

Sentence 1: The elephant hit the giraffe.

Sentence 2: The **lion** hit the elephant.

The critical manipulation was the accessibility change of the subject in the second sentence. Subjects (and agents in particular) tend to be highly accessible referents in a discourse (Givon, 1990). Thus, in the *Given-NonShift* trial, the target word 'lion' is highly accessible in both the first and second sentences, while in the *Given-Shift* condition, the target word shifts from a position of low accessibility (theme), to a position of high accessibility (agent) [2]. Finally, in the *New* condition, the target word has not been mentioned before, and thus also "shifts" to a prominent place in the discourse.

If children's use of acoustic prominence is sensitive to these shifts in discourse structure, we will see differences in the production of the target word across conditions. One possibility is that acoustic prominence will convey less articulated discourse representations in child speech. For example, children might distinguish the simpler, possibly more salient contrast between *given* and *new* information, but not more fine-grained accessibility contrasts. If so, we predict more acoustic prominence on the target word in the *New* condition (2c) than in either of the *Given* conditions (2a and 2b). On the other hand, if children use acoustic prominence in the same discourse situations as adults, they will also use more acoustic prominence on the target word in the *Given-Shift* condition (3b) than in the *Given-NonShift* (3a) condition.

METHODS AND PROCEDURE

Participants

28 children between the ages of 3 years 5 months and 4 years 9 months were tested individually in the presence of their parents. Each child came in to the lab for a 30–40 minute session which began with free play, intended to encourage the child to be comfortable speaking to the experimenter. Children received a small toy at the end of the experiment.

Procedure

In order to elicit the sentences, we showed children short video scenes displaying pairs of hand puppets performing actions. In all of the displays, one puppet was the agent of an action of which the other was the theme. There were three actions (hit, kiss, hug) and five puppets performing the actions. Each trial consisted of the child viewing two scenes (involving the same action) one after the other. After viewing the first scene they were asked to describe it

---

[2]An alternative possibility is that it is the *change* in accessibility, rather than the increase in accessibility, which leads to prominence (see Terken & Hirschberg, 1994). The current work does not differentiate between these two possibilities.

(e.g. *"What happened?"*). After the second scene, they were asked to describe the scene in a manner intended to convey a continuing discourse (e.g. *"And then what happened?"*). Due to the interactive nature of the experiment, there was some variability in the precise form of the questions. Children also occasionally produced the target sentence before the experimenter had the opportunity to prompt them. Children sometimes needed some coaxing to produce a suitable utterance. Where children did not spontaneously produce a sentence of the correct form, the experimenter would encourage them to produce the required sentence. (The experiment was designed such that the child could push a button to view the video as many times as (s)he wished). If the child did not finally produce the required sentence, the experimenter provided the sentence, using a neutral prosody with neither noun receiving emphasis, and asked the child to copy it. Trials in which this occurred for the target utterance were excluded from the analysis. We also excluded: trials in which the child or experimenter referred to any of the subject, object or verb of the target sentence before the target sentence was uttered; trials in which the child repeated or replaced one of the words in uttering the target sentence (e.g. *The lion --- no the bee hit the giraffe);* trials in which the target sentence had an unusual form which made it difficult to identify the subject and object nouns for analysis (e.g. *'it is putting the elephants eye on the giraffes stomach').* The decision as to whether a trial should be included was made by a coder blind to condition, who read a transcript of what occurred in each trial after the context sentence had been uttered. The trials were constructed such that each of the three nouns occurred as the target word *once* for each condition, making 9 trials in all (3 target nouns * 3 conditions – see Table 1). Trial order was randomized across participants. Participants took a short break between trials.

The session was recorded to a computer and analyzed using Praat acoustic analysis software (Boersma & Weenink, 2005)[3]. A research assistant labeled the word boundaries for subjects and objects in the target utterances and a script was used to automatically extract the acoustic measures. Three types of data were collected for both subjects and objects: (a) the *maximum pitch* (F0) reached at any point in the word (b) the overall *intensity* of the entire word and (c) the word's *duration*. (Since children did not maintain a constant distance from the microphone, we anticipated that our measure of intensity would be highly variable).

For each child, we calculated the average for each of these measures in each condition. Children with missing cells were excluded. Our criteria excluded all of the data from 14 out of the 28 child participants.

RESULTS

The critical data are the maximum pitch, the intensity and the duration of the target words (i.e. the subject nouns of the target utterances) in each condition. These can be seen in the top portion of Table 2. The maximum pitch is significantly higher in the *Given-Shift* and *New* conditions than in the *Given-NonShift* condition (*Given-Shift* v *Given-NonShift* $t=1.81$, $p< 0.05$, $df=13$, *New* v *Given-NonShift* $t=2.55$, $p<0.05$, $df=13$). There is no difference between the *Given-Shift* and *New* conditions ($t=0.77$, $p>0.05$, $df=13$). Following the same pattern, intensity is significantly greater in the *Given-Shift* and *New* conditions than in the *Given-NonShift* condition (*Given-Shift* v *Given-NonShift* $t=1.87$, $p< 0.05$, $df=13$, *New* v *Given-NonShift* $t=2.49$,

---

[3]We also conducted a subjective analysis of prosody. This was carried out by two coders, blind to experimental condition, who coded each target utterance as 'subject-accented' or 'subject-de-accented' using a variant of the ToBI coding system. The vast majority of utterances were coded as accented (*New* 90%; *Given-Shift* 96%; *Given-NonShift* 90%) and there were no significant differences across condition. This study thus adds to a growing body of research which finds differences in acoustic prominence across different discourse conditions *without* concurrent differences in the rate of accenting (Watson, Arnold, & Tanenhaus, 2005; Bard & Aylett, 1999), and which challenges the notion that acoustic emphasis can be described in terms of a simple categorical distinction between accented and de-accented words. The relationship between acoustic prominence and the subjective perception of accenting will be an important topic for future research, but is beyond the scope of the current study.

$p<0.05$, $df=13$). Again, there is no difference between the *Given-Shift* and *New* conditions ($t=0.50$, $p>0.05$, $df=13$). Recall, that the measurement of intensity was potentially influenced by the fact that children did not maintain a constant distance from the microphone. However, the close correlation with pitch suggests that this factor was not critical. There were no significant differences between any of the conditions for measurements of duration.

The target noun data thus demonstrate that this word is produced with greater pitch and intensity in the *Given-Shift* and *New* conditions than in the *Given-NonShift* condition. This result is in line with the adult data reported by Watson et al. (2005) and suggests that children emphasize words which have shifted in discourse accessibility. However a potential objection is that the properties of increased pitch and intensity could theoretically apply across a child's entire utterance, rather than specifically to the target noun. To explore this possibility, we also performed equivalent analyses of acoustic measures of the object noun in each sentence. These data are also shown in Table 2. It can be seen that these data do not pattern with the data collected from subject nouns. (In particular, note that object nouns were actually produced with lower pitch and less intensity in the *Given-Shift* than in the *Given-NonShift* condition). There were no significant differences between any of the conditions for any of the three measures for these data. The hypothesis that entire utterances are emphasized in the *Given-Shift* and *New* conditions is thus unsupported.

To further test the extent to which subject nouns receive greater comparative emphasis in *New* and *Given-Shift* conditions, we also calculated the difference in the pitch and intensity between subject and object nouns in each utterance. This provides a measure of the degree to which the target word "stood out" from other words in the sentence. The average for each condition is also shown in Table 2. Although subject nouns were produced with overall greater pitch and intensity than object nouns across conditions, the difference in intensity was found to be significantly greater in the *Given-Shift* and *New* conditions than in the *Given-NonShift* condition (*Given-Shift* v *Given-NonShift* $t=1.95$, $p<0.05$, $df=13$, New v *Given-NonShift* $t=2.11$, $p<0.05$, $df=13$), and the difference in pitch was found to be significantly greater in the *Given-Shift* than in the *Given-NonShift* condition (*Given-Shift* v *Given-NonShift* $t=2.43$, $p<0.05$, $df=13$). There was no reliable difference in the increase in pitch for the New condition than the *Given-Shift* condition ($t=0.05$, $p>0.05$, $df=13$). While the reasons for this lack of difference are currently unclear, overall these data support the hypothesis that our child participants use increased pitch and intensity specifically on the target nouns in the *Given-Shift* and *New* conditions.

DISCUSSION

The current work constitutes an important first step in investigating accent placement by young children. Four-year olds were found to use greater pitch and intensity on subject nouns that had shifted in discourse accessibility and less pitch and intensity on subject nouns that had not shifted.

One unexpected aspect of these data was that although emphasis was realized in differences in both pitch and intensity, there were no corresponding changes in duration. This is in contrast to Watson et al. (2005) who report increases in all three measures of acoustic prominence. However, even in adults, the extent to which the contribution of pitch, intensity and duration contribute to acoustic prominence is not well understood. Recent work suggests that these factors may be tied to differing functions in the discourse. For example, Watson et al. (in press) found that pitch and duration were tied to the predictability of information while intensity was tied to importance of information. Although the reason why children do not use duration to signal acoustic prominence in the current study is unclear, for our current purposes, the

critical result is that children's use of pitch and intensity correlates with differences in accessibility status.

We suggest that the different treatment of target nouns across the conditions reflects children's ability to track shifts in discourse accessibility. Note that the effects cannot be attributed to priming associated with simple lexical repetition (a phenomenon well documented in the literature, see Bard & Aylett, 1999). That is, the lower pitch and intensity in the *Given-NonShift* condition cannot merely be a result of previous mention since the critical word in the *Given-Shift* condition was also previously mentioned, (and in fact was mentioned more recently - a factor which would actually be expected to boost lexical priming). Whilst we acknowledge that it is in theory possible that lexical priming could be boosted by the particular conditions seen in the *Given-NonShift* condition, specifically the fact that when the lexical item is repeated it occurs in the same syntactic position and with the same thematic role, we point out that such an account, which could equally apply to the equivalent adult data, could not explain non-prosodic discourse phenomena, such as appropriate pronoun use. Since previous research has found that children are sensitive to discourse accessibility in pronoun production, we consider that the most plausible account of the current findings is that children's use of acoustic prominence reflects shifts in discourse accessibility. However, we acknowledge that our findings, and the adult data in Watson et al. (2005), are consistent both with an account in which the speaker explicitly represents the discourse, and with an account in which accessibility reflects a more passive aspect of processing. The extent to which prominence marking is listener or speaker centered is an important issue for future research.

Further research is also necessary in order to understand exactly how fine-grained the relationships between discourse shifts and acoustic prominence are. Watson et al. (2005) found what looked like a gradient relationship between prominence and accessibility shift: the greater the shift in accessibility, the greater the degree of acoustic prominence. Whether this also holds for child speech is an open question. A second question is whether children can *understand* accented information to refer to information that is not accessible in the discourse. The work here represents a first step in answering these questions.

Acknowledgements

References

Akhtar N, Carpenter M, Tomasello M. The role of discourse novelty in early word learning. Child Development 1996;67(2):635–645.

Arnold, JE.; Novick, JM.; Brown-Schmidt, S.; Eisenband, JG.; Trueswell, JC. Knowing the difference between girls and boys: The use of gender during on-line pronoun comprehension in young children. In: Do, AH-J.; Domínguez`, L.; Johansen, A., editors. Proceedings of the 25th annual Boston University Conference on Language Development; Cascadilla Press; 2001. p. 59-69.

Avrutin, S. Development of the syntax–discourse interface. Kluwer: Dordrecht; 1999.

Ariel, M. Accessing noun-phrase antecedents. London: New York: Routledge; 1990.

Bard, EG.; Aylett, MP. ICPHS-99. San Francisco, CA: 1999. The dissociation of deaccenting, givenness, and syntactic role in spontaneous speech.

Boersma, P.; Weenink, D. Praat: doing phonetics by computer (Version 4.3.36) [Computer program]. 2005 [Retrieved December 11, 2005]. from http://www.praat.org/

Bolinger, DLM. Intonation and its parts: Melody in spoken English. Stanford, Calif.: Stanford University Press; 1986.

Brown, G. Prosodic structure and the given/new distinction. In: Cutler, A.; Ladd, DR., editors. Prosody: Models and measurement. Berlin; New York: Springer-Verlag; 1983. p. 67-77.

Brown, RW. A first language: The early stages. Cambridge, Mass.: Harvard University Press; 1973.

Campbell AL, Brooks P, Tomasello M. Factors affecting young children's use of pronouns as referring expressions. Journal of Speech, Language, and Hearing Research : JSLHR 2000;43(6):1337–1349.

Chafe WL. Language and consciousness. Language 1974;50(1):111–133.

Cutler A, Dahan D, van Donselaar W. Prosody in the comprehension of spoken language: A literature review. Language and Speech 1997;40(Pt 2):141–201. [PubMed: 9509577]

Dahan D, Tanenhaus MK, Chambers CG. Accent and reference resolution in spoken-language comprehension. Journal of Memory and Language 2002;47(2):292–314.

Givon, T. Syntax: A Functional Typological Introduction. Amsterdam: John Benjamins; 1990.

Grosz BJ, Sidner CL. Attention, intentions, and the structure of discourse. Comput.Linguist 1986;12(3): 175–204.

Gundel JK, Hedberg N, Zacharski R. Cognitive status and the form of referring expressions in discourse. Language 1993;69(2):274–307.

Hickmann M, Hendriks H. Cohesion and anaphora in children's narratives: A comparison of English, French German, and Mandarin Chinese. Journal of Child Language 1999;26(2):419–452. [PubMed: 11706473]

Ladd, DR. Intonational Phonology. Cambridge: Cambridge University Press; 1996.

Perner J, Leekam SR. Belief and quantity: Three-year olds' adaptation to listener's knowledge. Journal of Child Language 1986;13(2):305–315. [PubMed: 3745334]

Prince, E. Toward a taxonomy of given-new information. In: Cole, P., editor. Radical pragmatics. New York, NY: Academic Press; 1981. p. 223-255.

Snow D, Balog HL. Do children produce the melody before the words? A review of developmental intonation research. Lingua 2002;112(12):1025–1058.

Song H, Fisher C. Who's "she"? Discourse prominence influences preschoolers' comprehension of pronouns. Journal of Memory and Language 2005;52(1):29–57.

Tomasello M, Call J, Hare B. Chimpanzees understand psychological states - the question is which ones and to what extent. Trends in Cognitive Sciences 2003;7(4):153–156. [PubMed: 12691762]

Watson, D.; Arnold, JE.; Tanenhaus, MK. Not just given and new: The effects of discourse and task based constraints on acoustic prominence; Poster presented at the 2005 CUNY human sentence processing conference; Tucson, AZ. 2005.

Watson D, Arnold JE, Tanenhaus MK. Tic Tac TOE: Effects of predictability and importance on acoustic prominence in language production. Cognition. (in press)

Table 1

Sentences for each of the 9 trials (target word in bold).

	Context Sentence	Target Sentence
New	The bee hit the ladybird.	The **giraffe** hit the lion.
	The elephant hugged the giraffe.	The **lion** hugged the bee.
	The bee kissed the giraffe.	The **ladybird** kissed the elephant.
Given-Shift	The elephant hugged the lion.	The **lion** hugged the bee.
	The bee kissed the ladybird.	The **ladybird** kissed the elephant.
	The ladybird hit the giraffe.	The **giraffe** hit the lion.
Given-NonShift	The ladybird kissed the giraffe.	The **ladybird** kissed the elephant.
	The giraffe hit the ladybird.	The **giraffe** hit the lion.
	The lion hugged the giraffe.	The **lion** hugged the bee.

Table 2

Acoustic Data. Data for the target subject nouns (top) with comparison data from object nouns and difference scores. (Standard error in parentheses).

		new	shift	nonshift
SUBJECT (TARGET) NOUN	maximum pitch within word (Hz)	312.64 (12.38)	323.93 (17.87)	302.37 (13.72)
	intensity across word (dB)	66.51 (1.72)	67.08 (1.69)	65.13 (1.69)
	duration of word (secs)	0.63 (0.05)	0.64 (0.04)	0.66 (0.05)

		new	shift	nonshift
OBJECT NOUN	maximum pitch within word (Hz)	296.27 (14.99)	279.08 (12.72)	286.63 (9.44)
	intensity across word (dB)	62.83 (1.41)	61.81 (1.62)	63.06 (1.59)
	duration of word (secs)	0.77 (0.05)	0.8 (0.06)	0.83 (0.03)

		new	shift	nonshift
DIFFERENCE: SUBJ-OBJ NOUN	maximum pitch within word (Hz)	16.37 (11.84)	44.85 (13.04)	15.74 (9.51)
	intensity across word (dB)	3.68 (0.81)	5.27 (1.36)	2.07 (0.43)
	duration of word (secs)	−0.14 (0.07)	−0.16 (0.06)	−0.17 (0.05)