Vowel Duration and Pitch Contour as Contenders for Infant Attention

Christine Kitamura and Jennifer Lorenzo

MARCS Auditory Laboratories
University of Western Sydney,
Australia
c.kitamura@uws.edu.au

Abstract

This study investigates infant preferences for the acoustic properties of vowels. For this purpose F_0 contour shape (bell or steady-state) and duration (normal or long) of the corner vowels /i/ and /u/ were manipulated in the words, ‘bee’, ‘shoe’, ‘goose’, and ‘sheep’. Four experiments used a two-choice preference method. The results show that when the words were of normal duration, infants prefer bell-shaped over steady-state contours in Experiment 1, but when both words were of long duration in Experiment 2, they show no preference. In Experiment 3 and 4 comparing long to normal duration vowels infants show a preference for long over normal duration words but only when the contours were bell shaped. It is concluded that the shape of the vowel is more important to infant attention than increased duration.

1. Introduction

The advantages of infant-directed speech (IDS) are multifaceted - this primitive mode of communication appears to be a powerful way of maintaining infant attention, communicating emotion, and of teaching infants the protocols of social interaction (Kitamura, Thanavisuth, Luksaneeyanawin, & Burnham, 2002). When compared to adult-directed speech (ADS), IDS has high affective salience, slower speech rate, higher pitch, greater pitch modulation, more distinctive pitch contours, and hyperarticulated vowels (Kitamura & Burnham, 2003; Fernald & Simon, 1984; Andruski & Kuhl, 1996) and it expresses meaning to the infant (Kitamura & Burnham, 2003). It is well established that infants prefer IDS to ADS (Fernald, 1985; Werker & McLeod, 1989; Pegg, Werker & McLeod, 1992; Panneton Cooper & Aslin, 1990) and that these preferences are based on its inherent affective salience (Kitamura & Burnham, 1998) carried by features such as slower tempo (Panneton Cooper, Kitamura, Mattock & Burnham, under revision) and variations in fundamental frequency (F_0) (Katz, Cohn & Moore, 1996).

While infants have definite attentional and affective preference for IDS, studies are beginning to suggest that, over and above this, IDS has a linguistic, didactic function. Indirect evidence comes from analyses of the features of IDS, especially its hyperarticulated vowels, which interestingly, Burnham, Kitamura and Vollmer-Conna. (2002) have shown to be specific to IDS as they are not found in similar speech styles such as speech directed to pets. One of the few studies showing evidence of a didactic function of IDS found that 1- to 4-month-old infants can discriminate /marana/ from /malana/ only when the medial syllables /ra/ and /la/ are synthesized using stress characteristics typical of IDS, such as increases in contour, duration and intensity (Karzon, 1985). More recently, infant discrimination ability has been linked to mother’s degree of hyperarticulation (Liu, Kuhl & Tsao, 2003).

It seems that the information contained in vowels is critical to infant attention. Cutler (1994) suggests that young infants have a ‘periodicity bias’, responding preferentially to the periodic information in vowels over the broadband information in consonants. Vowels in IDS differ from those in adult-directed speech – they are hyperarticulated, have higher and more variable pitch, increased positive affect, and longer duration (Burnham et al., 2002). While IDS has formant values that expand its vowel space, its vowels have other acoustic features such as elevated and more variable F_0, and longer duration, which may act to facilitate discrimination of words and speech sounds. Trainor & Desjardins (2002) found that F_0 contour shape (falling vs. steady state) but not mean F_0 (high vs. low) facilitates infant discrimination of vowel categories. The experiments reported here extend the findings of

Proceedings of the 10th Australian International Conference on Speech Science & Technology
Macquarie University, Sydney, December 8 to 10, 2004.

Accepted after abstract only review

Trainor and Desjardins (2002) but report infant *attention* to two acoustic properties of vowels: F_0 contour shape and vowel duration. Increased vowel duration is included because it has been found that slowing the speech rate of IDS increases infant attention (Panneton Cooper et al. under revision). The aim of this study is to investigate whether infants' attention to words is governed by F_0 contour shape or increased vowel duration, or whether there is an interaction between these features.

2. Method and Results

Using a two-choice preference procedure, 4 groups of infants were given the opportunity to listen to familiar word lists in one of four experiments:

1. Normal duration (bell versus steady-state contours)
2. Long duration (bell vs steady-state contours)
3. Bell contours (normal vs long duration)
4. Steady-state contours (normal vs long duration).

Thus the shape of the F_0 contour could be bell shaped with high F_0 range or be a steady-state F_0 contour with low F_0 range and/or it could be of normal or long duration. Each speech sample was matched with a visual target consisting of a multi-colored target.

2.1 General Method

Infants were tested seated on their parent's lap facing two video monitors with separate speakers. Each set of vowel exemplars was paired with the same visual target, one presented to the infants' left and the other to their right. Thus if infants fixated the left visual target they heard one set of vowel exemplars and if they fixated the right visual target they heard the other set. Infants were first familiarised for 20-seconds with each of the stimulus sets, which played irrespective of the infant's looking behaviour. In the test phase, there were six 20-second trials in which presentation of the speech stimulus was contingent on the infants' direction of fixation. Side of presentation was counterbalanced so that half the infants heard one set of vowel exemplars on the left side and the other set of vowel exemplars on the right side and this was reversed for the other half of the. participants. In addition, order of presentation in the familiarisation phase was counterbalanced: half the infants began trials on the left side and the other half on the right side. During familiarisation and testing the infants' head and eye movements were recorded by video camera and observed on a video monitor in an adjacent room. The observer viewed the infant's image and used a left-right toggle switch to turn on the left or right speech stimuli depending on the infants' looking direction.

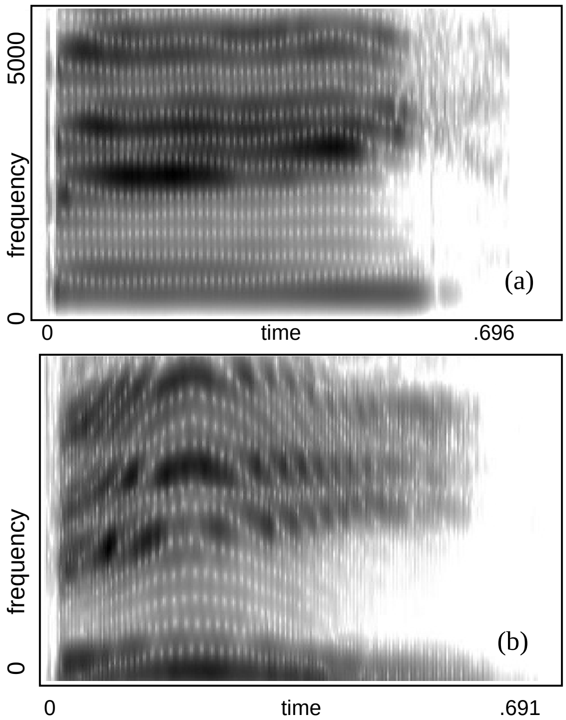

Figure 1: Spectrograms of the word 'bee'. An example of a steady-state contour is in the top panel (a), and bell-shaped contour is in the lower panel (b).

Table 1: Measures of starting F_0, vowel duration, and F_0 range of word exemplars used in Experiments 1-4.

	Steady state			Bell shape		
exemplar	**F_0**	**dur**	**range**	**F_0**	**dur**	**range**
goose_1	322	0.68	34	271	0.71	313
goose_2	287	0.73	76	315	0.70	326
goose_3	287	0.75	76	343	0.76	321
goose_4	289	0.76	30	320	0.83	297
Mean	**296**	**0.73**	**54**	**312**	**0.75**	**314**
bee_1	298	0.59	27	238	0.59	377
bee_2	343	0.56	37	365	0.55	354
bee_3	302	0.61	98	290	0.77	224
bee_4	302	0.63	98	340	0.67	451
Mean	**311**	**0.60**	**65**	**308**	**0.65**	**352**
shoe_1	289	0.80	32	342	0.79	295
shoe_2	356	0.73	33	319	0.77	119
shoe_3	322	0.65	39	306	0.90	217
shoe_4	323	0.70	41	261	0.61	380
Mean	**323**	**0.72**	**36**	**307**	**0.77**	**253**
sheep_1	314	0.67	48	311	0.74	268
sheep_2	337	0.83	57	354	0.74	268
sheep_3	312	0.77	48	316	0.76	237
sheep_4	313	0.80	57	347	0.81	287
Mean	**319**	**0.77**	**53**	**332**	**0.76**	**265**

Accepted after abstract only review

For the vowel stimuli, a female speaker, using an IDS register, produced multiple tokens of familiar words containing the corner vowels /i/ in the words 'bee' and 'sheep,' and /u/ in the words 'shoe' and 'goose' (MacArthur CDI; Fenson et al., 1993). The female speaker produced the word tokens with and without the distinctive F_0 contour (high F_0 range and low F_0 range, respectively shown in Figure 1). The word stimuli selection was based on matching the mean F_0 and duration of the vowels in the four sets of words to be used as exemplars of bell or steady-state contours. To ensure infants were not making choices based on differences in perceived pitch, the level of F_0 of the steady-state exemplars was matched to the starting F_0 period of the bell-shaped contours. The selected word stimuli, both bell and steady-state contours, were also matched for vowel duration. The measures are shown in Table 1. The order of the stimuli was shoe, bee, goose, sheep. For each of the four exemplars of each word, the duration of the vowels was doubled using SoundEdit, and the 16 word stimuli appended and looped with a 300ms gap between words.

2.2 Experiment 1: Normal Duration (Bell versus Steady-state Contours)

In this experiment infants were presented with familiar words in which the mean F_0, and the duration of the vowels in those words was equated but the shape of the F_0 contour manipulated so that on one side the vowel exemplars available to the infant had a bell-shaped F_0 contour with high F_0 range and on the other side, the vowel exemplars were steady state F_0 contours with low F_0 range. Twenty-one 6-month-old infants were tested (mean=25.84 weeks; range=24.3-27.6 weeks; 9 females, 12 males). It is expected that infants will prefer bell to steady-state contours.

2.2.1 Results

The data were analysed in 2 x 2 x (2) ANOVA with side of presentation (left or right) and order of presentation in the familiarization trials (Bell first or Steady-state first) as between-subjects factors, and looking duration (Bell or Steady-state) as the within-subjects factor. The results show no significant main effects or interactions for side or order counterbalancing but there was a significant main effect for looking duration showing that infants looked longer to bell-shaped contours than steady-state contours($F(1,19)=7.42$, $p=.013$, $\eta^2=.281$).

2.3 Experiment 2: Long duration (Bell versus Steady-state F_0 Contours

Experiment 2 replicated the stimulus conditions of experiment 1 except the duration of the bell and steady-state contours was lengthened to twice normal duration. Twenty six-month-old infants were tested (mean age = 26.74, range=23.71-28.71; 10 males and 10 females).

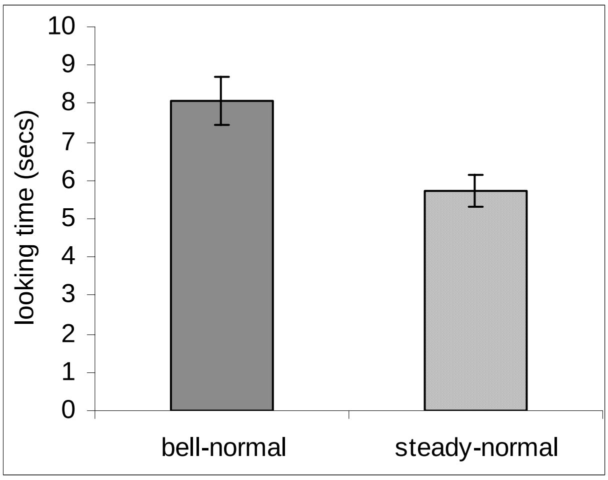

Figure 2: Looking times to bell-normal and steady-state normal contours in Experiment 1.

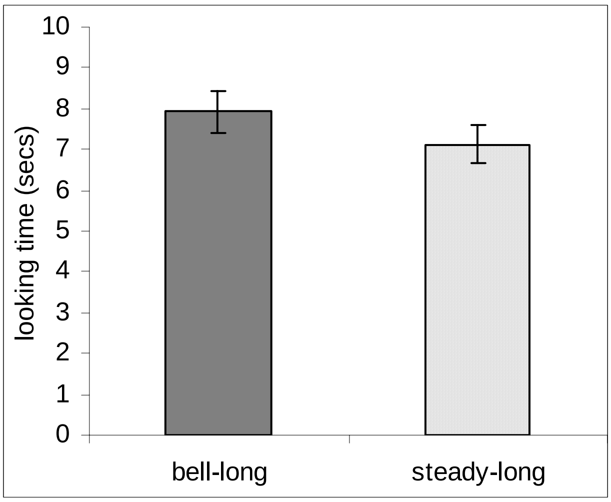

Figure 3: Looking times to bell-long and steady-long contours in Experiment 2.

2.3.1 Results

The data were analysed in a 2 x 2 x (2) ANOVA with order and side of presentation as between-subjects' variables and looking duration (bell or steady-state long contours) as the within-subjects variable. There were no significant results for order, side or looking duration ($p>.05$). As shown in Figure 2 infants found long bell and long steady-state contours equally interesting.

2.4 Experiment 3: Bell-shaped Contours (Normal versus Long duration)

In Experiment 3, bell-shaped contours were presented to the infants: the duration was normal on one side and twice the length on the other side. Twenty-one six-month-old infants were tested.

Proceedings of the 10th Australian International Conference on Speech Science & Technology
Macquarie University, Sydney, December 8 to 10, 2004.

Accepted after abstract only review

2.4.1 Results

The data were analysed in a 2 x 2 x (2) ANOVA with side and order as between subjects' variables and looking time as the within subjects' variable. The results show no significant main effects or interactions for order or side of presentation. However, there was a significant main effect for looking time, infants looked longer to bell contours with increased duration over normal length bell-shaped contours, $F(1,19)=12.03$, $p=.003$, $\eta^2=.388$.

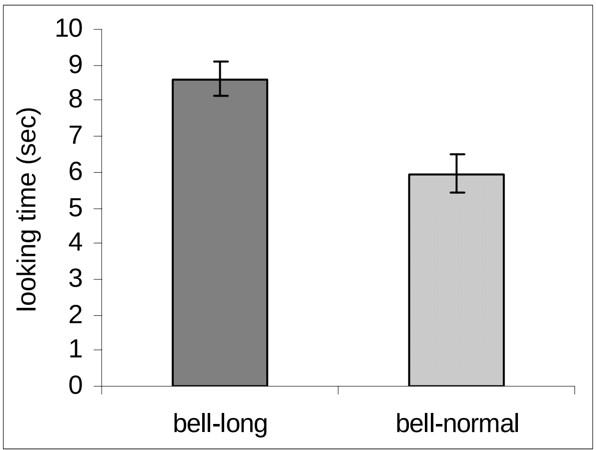

Figure 4: Looking times to bell-long and bell normal contours in Experiment 3

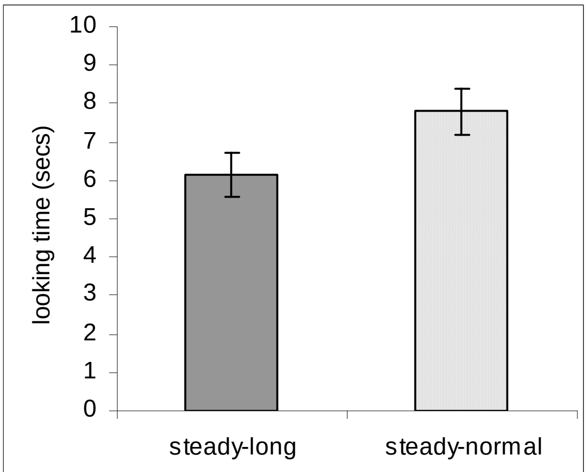

Figure 5: Looking times to steady-long and steady-normal contours in Experiment 4

2.5 Experiment 4: Steady-state Contours (Normal versus Long Duration).

In Experiment 4, steady-state contours were presented to infants, they heard normal duration steady-state contours on one side and steady-state contours doubled in duration on the other side. Twenty six-month-old infants were tested (mean age=26.51 weeks; range=24.29-28.71 weeks; 12 females and 8 males).

2.5.1 Results

The data were analysed in a 2 x 2 x (2) ANOVA with order and side of presentation as between subjects' variables and looking duration (Normal or Long contours) as the within subjects variable. There were no significant results for order, side or looking duration ($p>.05$). As shown in Figure 4 infants found normal length steady-state contours as interesting as steady-state contours that had been doubled in duration. Interesting the trend is for infants to prefer steady-state normal contours.

3. Discussion

The results show that 6-month-old infants (i) prefer bell-shaped to steady-state contours but only when both are of normal duration (Experiment 1) not when shape is confounded by increased duration (Experiment 2), and (ii) prefer listening to long over normal duration contours but only when both vowel stimuli are bell-shaped (Experiment 3) not when they are steady state (Experiment 4). At first glance, the data support the hypothesis that infant attention is driven by an interaction between contour shape and duration. However, it can be argued that increased F_0 variability or contour shape is the more influential of the two variables. In fact, infants show less preference for long steady-state vowels than those of normal length in experiment 4, and it seems increased duration interferes with infant preferences in experiment 2. Durational cues only increase attention when they are combined with increased F_0 variability (Experiment 3). The results also provide support for Trainor and Desjardins (2002) study showing that contour shape but not increased mean F_0 facilitates vowel discrimination. In fact, they suggest that high mean F_0 may even be detrimental to vowel perception. Similarly in this study infants show no preference for steady-state vowels, and although mean F_0 was not a variable here, vowels in the steady-state conditions had high mean F_0 and low F_0 variability. So it appears that infant sensitivity is to the shape of the contour more than its durational qualities.

While revealing 6-month-old infants have definite preferences for particular acoustic features of vowels, we need to be careful of applying this to other infant ages. Previous research with connected speech shows 4.5-month-old infants pay attention to speech with slowed tempo and to speech with high vocal affect (*even when both speech samples are slowed*) and show no preference when vocal affect and duration are pitted against each other. This indicates that both vocal affect and duration are equally important to the infant at this age, and indeed duration may cue positive vocal affect to young infants. Eight-month-old infants, on the other hand, do not show a preference for slowed speech or high affect speech as 4-month-old infants do. Here however, although we have infants listening to lists of words rather than connected speech, there appears to be remnants of the preference for increased duration at 6

Accepted after abstract only review

months because we still find infants preferring long to normal length bell-shaped contours.

At six months of age, infant pay attention to vowels with high F_0 range, presumably because movement in the voiced section of the speech signal holds infant attention. Whether six-month-olds are interpreting F_0 movement as an expression of emotion is not shown here. Certainly we know that 6-month-old preferences for infant directed connected speech are based on its affective salience and not its F_0 characteristics per se (Kitamura & Burnham, 1998) although F_0 movement signals heightened emotion in speech (Scherer, 1986) to infant or adults alike. In addition, F_0 movement might make the discrimination of vowels easier because the harmonics associated with the F_0 contour moving through the resonant frequencies of the formants might aid their location (Trainor & Desjardins, 2002). Thus infant may prefer bell-shaped contours either because (1) they pay more attention to more complex auditory stimuli (Columbo, 1986) or (2) they are interpreting movement in the vowels as an expression of positive emotion. Either way this would advantage the discrimination and acquisition of vowel categories.

4. Acknowledgements

This research was funded by a Research Grant from the University of Western Sydney Grant (20211/80401).

5. References

Andruski, J., & Kuhl, P. K. (1996). The acoustic structure of vowels in mothers' speech to infants and children. In *Proceedings of the Fourth International Conference on Spoken Language Processing* (pp. 1545-1548). Philadelphia.

Burnham, D., Kitamura, C., Vollmer-Conna, U. (2002). What's new pussycat: On talking to animals and babies. *Science, 296*, 1435.

Colombo, J. (1986). Spectral complexity and infant attention. *The Journal of Genetic Psychology, 146*, 519-526.

Cutler, A. (1994). Segmentation problems, rhythmic solutions. *Lingua, 92*, 81-104.

Fenson, L., Dale, P.S., Reznick, J.S, Thal, D., Bates, E., Hartung, J.P., Pethick, S., & Reilly, J.S. (1993). *The MacArthur Communicative Development Inventories: User's guide and technical manual*. San Diego: Singular Publishing Group.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development, 8*, 181-195.

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mother's speech to newborns. *Developmental Psychology, 20*, 104-113.

Karzon, R. (1985). Discrimination of polysyllabic sequences by one to four month old infants. *Journal of Experimental Child Psychology, 39*, 326-342.

Katz, G. S., Cohn, J. F., & Moore, C. A. (1996). A combination of vocal F_0 dynamic and summary features discriminates between three pragmatic categories of infant-direct speech. *Child Development, 67*, 205-217.

Kitamura, C., & Burnham, D. (1998). The infant's response to vocal affect in maternal speech. In C. Rovee-Collier (Ed). *Advances in Infancy Research, 12*, 221-236.

Kitamura, C., & Burnham, D. (2003). Pitch and communicative intent in mothers' speech: adjustments for age and sex in the first year. *Infancy, 4*, 85-110.

Kitamura, C., Thanavisuth C., Luksaneeyanawin, S., & Burnham, D. (2002). Universality and specificity in infant-directed speech: Pitch modifications as a function of infant age and sex in a tonal and non-tonal language. *Infant Behavior and Development, 24*, 372-392.

Liu, H., Kuhl, P. K., & Tsao, F. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science, 6*, F1-F10.

Panneton Cooper, R., Kitamura, C., Mattock, K., & Burnham, D. (under revision). Developmental differences in infants attention to affective properties of infant-directed speech. *Infancy*.

Panneton Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development, 61*, 1584-1595.

Pegg, J. E., Werker, J. F., & McLeod, P. J. (1992). Preference for infant-directed over adult-directed speech: Evidence from 7-week-old infants. *Infant Behavior and Development, 15*, 325-345.

Scherer, K. R. (1986). Vocal affect expression: A review and model for future research. *Psychological Bulletin, 99*, 143-165.

Trainor, L. J., & Desjardins, R. N. (2002). Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. *Psychonomic Bulletin and Review, 9*, 335-340.

Werker, J. F., & McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional affective responsiveness. *Canadian Journal of Psychology, 43*, 230-246.

Accepted after abstract only review