

INTRINSIC F0 IN THE BABBLING OF MANDARIN-LEARNING INFANTS

D. H. Whalen, Andrea G. Levitt¹, Pai-Ling Hsiao², Julia Irwin² and Winifred McGowan
Haskins Laboratories (¹also Wellesley College) (²also University of Connecticut)

ABSTRACT

Intrinsic F0 (IF0) is the tendency for high vowels to have a higher F0 than low vowels. We previously found IF0 in babbling of French- and English-learning infants, suggesting an automatic effect. Here, we find IF0 in four 12-month-old Mandarin babblers, even though they are learning a tone language. Thus it seems that IF0 is to be explained not as an enhancement of the vowel quality difference but rather as an automatic consequence of vowel formation.

INTRODUCTION

Languages differ on many dimensions, but other features are consistent across languages. One phonetic feature that has been found to accompany vowels is "intrinsic F0" or "intrinsic pitch" (IF0, from here on). This is the tendency for the high vowels, such as [i] and [u], to have a higher fundamental frequency than the low vowels, such as [a] and [æ]. IF0 was first noticed for German [1] and has since been found in every language that has been examined for it. In a previous survey, we found evidence of IF0 in 31 languages from 11 of the 29 major language families of the world, and no instances of a lack of IF0 [2].

There has been considerable debate about the mechanism responsible for this effect. The present work does not directly support any particular explanation, so the reader is directed to the various surveys of the explanations to be found elsewhere [3-7]. All of the prior explanations have assumed that IF0 is an automatic consequence of articulation, probably due to the pull of the tongue on the laryngeal system, or to an acoustical interaction between F1 and F0. (Steele [8] argues that there must be a contribution of subglottal pressure.) Some authors suggest that only a combination of explanations can account for all of the facts.

However, more recently there have been proposals that IF0 is a deliberate manipulation of F0 that is introduced in

the signal to enhance the differences between vowel categories [9-11]. On this account, speakers try to accommodate their listeners by making F0 closer to F1 for high vowels and farther from F1 for low vowels. There is some evidence that listeners perceive vowel height not in terms of F1 by itself but by a difference between F1 and F0 [12]. The enhancement account, then, asserts that speakers have control over this aspect of the F0, and IF0 is simply a particularly useful enhancement.

There are two predictions that come from the enhancement account. First, there should be some population that chooses not to enhance its vowels in this way. Second, there should be some developmental change in the use of IF0 if it is an enhancement that needs to be learned.

The first prediction, of a language which chooses not to use IF0, has not been borne out. The survey of Whalen and Levitt [2] found no instances of languages which lacked IF0. The survey included languages with quite different vowel inventories, and still found no difference across languages, even though an enhancement would seem to be more useful in a crowded space than in a sparse one. It is never possible to prove that there is no language that exhibits a certain trait, since an example could be waiting to be discovered. For IF0, however, the sample of languages was broad and also included cases where one would expect to find no difference. A cogent example of such a language is one which uses F0 for lexically distinctive tones and, at the same time, has a small vowel inventory. (Enhancement by IF0 would seem to be most useful in crowded vowel spaces.) One such language is Mandarin, and Mandarin has in fact been shown to have IF0 [13].

The second prediction, that the IF0 effect should change during language development, has also not been supported. Whalen, Levitt, Hsiao and Smorodinsky [14] examined infants in

two language environments (French and English), at the ages of 6, 9 and 12 months. Despite the fact that these same infants showed significantly different use of F0 for intonation [15], they showed a typical IF0 effect that did not differ across language environment or across age. That paper also surveyed six studies with older children (6-11 years), and found no developmental trend at older ages either. If the IF0 effect is present from the beginning of linguistic production (and there is very little linguistic phonation before 6 months), it seems very unlikely that it is a learned enhancement.

An enhancement account might assume that infants are imitating IF0. It is true that every language the infant hears will show IF0 (since it is universal), and it is thus logically possible that the IF0 in babbling is imitative. However, it is not clear how the infant would know to extract this property of the signal, since the infant lacks vowel categories in the babbling stage. Furthermore, the speech directed to infants ("motherese") contains very large changes in F0 [16], which would make the extraction of the relation between vowel height and F0 that much more difficult. Finally, children learning a tone language would also hear each vowel at very different F0s, depending on the tone used with it. All of these factors make the task of detecting the IF0 extremely difficult for the child.

Nonetheless, if any population were to benefit from avoiding IF0, it would seem to be learners of a tone language. Tone is crucial for lexical distinctions, and it depends largely on F0, which is a phonetic dimension that seems to be under the infant's control earlier than segmental ones. Indeed, tonal categories seem to be mastered sooner than segmental categories [17]. Even if the IF0 contribution were a deliberate enhancement in tone languages, it would seem that the learner of a tone language would be most likely to use F0 just for emerging tone distinctions instead. In order to test this directly, we recorded the babbling of four Mandarin-learning infants and measured the F0s of the vowels to see whether the IF0 effect found for French- and English-learning

infants also appeared for these infants learning a tone language.

THE EXPERIMENT

We measured the F0 of all non-central vowels in the babbling of four Mandarin-learning infants.

Subjects

The infants were being raised as monolingual speakers of Mandarin. Infants were selected for the study only if both parents were native speakers of Mandarin Chinese. Most were from the Beijing area. The children were living in Storrs, Connecticut, while one or both parents attended the University of Connecticut. Most of these students planned to return to Mainland China after graduation and were therefore raising their children as monolingual speakers of Mandarin. One of the four had a monolingual Mandarin-speaking grandparent taking care of him for the duration of the recording sessions. Another had a grandparent visiting during part of the recording period.

Recordings

The infants were recorded in the home every other week for a session lasting 30-45 minutes. Recordings started at six or seven months of age and ended at 11-16 months. A Panasonic SV-3700 DAT tape recorder was used in conjunction with a Realistic wireless microphone. The microphone itself was sewn into a vest (concealed as the center of a flower) which the infant wore during the session. In this way, a relatively constant distance between the infant's mouth and the microphone could be maintained without restraining the child.

Analysis

The recordings were transferred to a VAX computer for analysis. The utterances were selected as being speech-like and separated from other sounds by 750 ms or more. All utterances were then transcribed by a native speaker of Mandarin. The symbols of the IPA were used, with the understanding that some of the utterances would be ambiguous at this level of detail.

For the present analysis, only the 12 month recordings were used. One subject (BX) returned to China in his

eleventh month, so his eleventh month recordings were used instead. We further restricted the analysis to non-central vowels; we also excluded /e/ (a very common vowel in these transcriptions) as a practical way of reducing the number of tokens to be analyzed without sacrificing the points of interest.

F0 was measured by delimiting ten pitch periods by hand in the acoustic waveform. This was performed with the program HADES, written at Haskins Laboratories [18]. We tried to take a measurement at a point 40% of the way into the syllable. In the best case, there were five periods on either side of that point, which would give us a single, average value for that stretch of speech. If that portion turned out to be unmeasurable, a stretch of ten periods as close to the 40% point and still within the syllable was found. These selection criteria resulted in 3155 vowel tokens that were measured. F0 values greater than 850 Hz were excluded to reduce the influence of occasional outlier, resulting in a final analysis of 3054 tokens. Of these, 2752 were transcribed as /e/.

RESULTS

As Table 1 shows, IF0 is present in the babbling of these four Mandarin-learning 12-month-olds. (Analyzing the results according to front/back as well as height was not possible because all subjects had gaps in their results that way.) The one negative difference (TZ) can be presumed artifactual, because of the small number of low vowels for this subject. Similarly, the large difference for BX also depends on a small number of low (and high) vowel tokens. The other two subjects show a difference of just the size we would expect based on our previous work with babbling. The overall difference is smaller than expected only because a large proportion of the low vowels happen to come from a speaker (EW) with a high overall F0.

DISCUSSION

Intrinsic F0 (IF0), which has been found in every language measured so far and in our previous study of babbling, has been found here in the babbling of Mandarin-learning infants as well. The size of the effect is of the same magnitude as in the earlier study of French- and English-learning infants.

Table 1. Average F0 for vowels of three heights (in Hz) for the four subjects and in a weighted average across subjects. Number of tokens is given below the F0 value.

Vowel Height	High	Med	Low	H-L
EW	404 (41)	369 (781)	348 (82)	56
TZ	328 (25)	293 (718)	336 (17)	-8
YL	332 (70)	311 (996)	279 (25)	53
BX	395 (10)	346 (274)	253 (15)	142
Mean	356	326	324	32

Thus even in a language that uses a sparse vowel space and lexical tones, infants exhibit IF0 in their own productions.

These results are incompatible with the notion that IF0 is a deliberate enhancement of the speech signal. That position assumes that the increase in F0 for high vowels helps to shift the effective F1 and thus enhance the vowel category differences. Infants learning Mandarin need to learn to produce the tone contours if they are to become successful speakers. Therefore, they have every reason to attend to the tonal aspects of F0 and to ignore, if possible, confounding factors such as IF0. If there was a population that would seem to benefit from ignoring this (potential) enhancement, it would appear to be the Mandarin learners. The fact that they do not is further evidence that IF0 is not an enhancement. Rather IF0 appears to be an automatic consequence of vowel production (from whatever source or combination of sources), even in infants.

ACKNOWLEDGMENT

This work was supported by NIH grant DC-00403 to Catherine T. Best and Haskins Laboratories.

REFERENCES

- Meyer, E.A. (1896-7), "Zur Tonbewegung des Vokals im gesprochenen und gesungenen einzelwort." *Phonetische Studien (Beiblatt zu der Zeitschrift Die Neuren Sprachen)*, vol. 10, pp. 1-21.
- Whalen, D.H. and A.G. Levitt. (in press), "The universality of intrinsic F0 of vowels." *Journal of Phonetics*, in press.
- DiCristo, A., D.J. Hirst, and Y. Nishinuma. (1979), "L'estimation de la F0 intrinsèque des voyelles: etude comparative." *Travaux de L'Institut de Phonétique D'Aix-en-Provence*, vol. 6, pp. 149-176.
- Shadle, C.H. (1985), "Intrinsic fundamental frequency of vowels in sentence context." *Journal of the Acoustical Society of America*, vol. 78, pp. 1562-1567.
- Fischer-Jørgensen, E. (1990), "Intrinsic F0 in tense and lax vowels with special reference to German." *Phonetica*, vol. 47, pp. 99-140.
- Sapir, S. (1989), "The intrinsic pitch of vowels: Theoretical, physiological and clinical considerations." *Journal of Voice*, vol. 3, pp. 44-51.
- Silverman, K.E.A. (1987), *The structure and processing of fundamental frequency contours*. Unpublished Ph.D. thesis, University of Cambridge.
- Steele, S.A. (1986), "Interaction of vowel F0 and prosody." *Phonetica*, vol. 43, pp. 92-105.
- Diehl, R.L. and K.R. Kluender. (1989), "On the objects of speech perception." *Ecological Psychology*, vol. 1, pp. 121-144.
- Diehl, R.L. (1991), "The role of phonetics within the study of language." *Phonetica*, vol. 48, pp. 120-134.
- Kingston, J. (1993), "The phonetics and phonology of perceptually motivated articulatory covariation." *Language and Speech*, vol. 35, pp. 99-113.
- Trautmüller, H. (1981), "Perceptual dimension of openness in vowels." *Journal of the Acoustical Society of America*, vol. 69, pp. 1465-1475.
- Shi, B. and J. Zhang. (1987), "Vowel intrinsic pitch in standard Chinese", in *Proceedings XIth International Congress of Phonetic Science*, pp. 142-145. Academy of Sciences of the Estonian SSR: Tallinn, Estonia.
- Whalen, D.H., A.G. Levitt, P.-L. Hsiao, and I. Smorodinsky. (1995), "Intrinsic F0 of vowels in the babbling of 6-, 9- and 12-month-old French- and English-learning infants." *Journal of the Acoustical Society of America*, vol. 97, pp. 2533-2539.
- Whalen, D.H., A.G. Levitt, and Q. Wang. (1991), "Intonational differences between the reduplicative babbling of French- and English-learning infants." *Journal of Child Language*, vol. 18, pp. 501-516.
- Fernald, A., et al. (1989), "A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants." *Journal of Child Language*, vol. 16, pp. 477-501.
- Clumeck, H. (1980), "The acquisition of tone", in *Child phonology, volume 1: Production*, G. Yeni-Komshian, J.F. Kavanagh, and C.A. Ferguson, Editor, pp. 257-275. Academic Press: New York.
- Rubin, P.E. (1995), "HADES: a case study of the development of a signal analysis system", in *Applied Speech Technology*, A. Syrdal, R. Bennett, and S. Greenspan, Editor, pp. 501-520. CRC Press: Boca Raton, FL.