Melodic complexity in infant language development[*]

Kuniya Nasukawa

Abstract

This paper employs the notion of melodic complexity in analysing the mismatch between perception and production in the acquisition of laryngeal-source contrasts in Japanese infants. Adopting the view that child phonology is governed by the same computational characteristics as adult phonology, it claims that the disparity is attributed to the result of immature performance systems which cannot decode complex structures in phonological representations.

1. Introduction

In contrast to Halle (1995), this paper takes the view that the domain of phonology is concerned solely with abstract entities and their relational properties, which means that phonological representations contain no instructions for the physical execution of speech. In such a model, I assume that some performance systems consist of a device for decoding phonological representations and mapping them on to physical entities.

This discussion considers how the relation between this decoding device and phonological representations is involved in phonological acquisition, focusing on the disparity between perception and production in child speech. It will be claimed that the disparity is attributed to immature reception and transmission facilities which cannot decode certain complex structures present in phonological representations at a very early stage in the developmental path.

Focusing on the acquisition of laryngeal-source contrasts in Japanese infants, this paper considers how those facilities develop diachronically in relation to phonological representations. First, §2 discusses the different levels of phonological representation and how these are relevant to the analysis of phonological acquisition. Next, in §3, within a cross-linguistic context I identify the laryngeal-source contrasts which are employed in Japanese. §4 provides data showing the diachronic development of speech production and perception in Japanese infants, both of which are based on my own research.

---

[*] In Marina Tzakosta, Claartje Levelt & Jeroen van de Weijer (eds.), *Developmental Paths in Phonological Acquisition*. Special issue of *Leiden Papers in Linguistics* 2.1 (2005), pp. 53-70. http://www.ulcl.leidenuniv.nl
ISSN 1574-4728.

After explaining the fundamental points of element-based feature theory in §5, §6 will discuss the relation between phonological representations and the reception and transmission facilities in the development of laryngeal-source contrasts in Japanese. Finally, my conclusions are given in §7.

2. Phonology and acquisition

2.1. Representations in phonology

In phonological studies based on the theory of Universal Grammar (UG), formal statements of dynamic and static regularities are generally considered from two viewpoints – their derivational and representational aspects. The derivational aspect is concerned with controlling the set of processes for mapping lexical phonological representations onto derived representations. On the other hand, the representational aspect involves expressing the formal relations between phonological categories, to which derivational processes apply. In earlier models of UG-based generative phonology, a lot of attention is paid to the relatively large set of processes undertaken during the course of derivation. By contrast, in more recent years the derivational aspect of phonological processing is greatly simplified and limited to the set of basic operations, and the research focus has shifted from the derivational towards the representational aspect.

The representational aspect of phonology accounts for the relational properties of phonological categories. It is widely acknowledged that phonological forms consist of two distinct types of representation: *prosody* comprises a hierarchy of domains formed by prosodic categories (e.g. skeletal positions, nuclei, or morae), while *melody* consists of melodic categories (e.g. features, elements or components) and their relational aspects. In both representational domains, the relations holding between formal categories are governed by principles (e.g. Onset Maximisation and Feature-Geometric organisation) provided by UG.

In the theory of UG, there is a basic premise that children are equipped with a set of innate prosodic and melodic categories and their relational properties governed by principles, which shape the direction of phonological acquisition. During language development, children acquire the language-specific aspects of grammar, which play only a peripheral role in UG (Fee 1995: 51-52).

2.2. The disparity between perception and production in child speech

It is widely acknowledged that there are disparities between perception (recognition) and production in the developmental path of language sounds among very young children. Specifically, it has been reported that perceptual ability is acquired much earlier than production ability. Infants have an innate perceptual ability to distinguish various aspects of utterances in a given language at a very early stage of acquisition (Jusczyk 1997: 86-87, Mehler *et al*

1996, Smith 2003). For example, infants can aptly discriminate minimal-pair contrasts before they can pronounce either of the forms concerned. This ability becomes fine-tuned during the course of acquisition, and loses some linguistically irrelevant differences. This developmental path leads us at least to the assumption that infants have access to the full set of universal melodic categories, such as features, which they then use as the basis for constructing phonological representations according to what they hear in the target language (cf. Smith 1973, Braine 1976 and Smolensky 1996). The production phase of infants also depends on such representations, but accurate forms are produced at a much later stage of development.

The disparity between perception and production is in principle based upon one of the following two perspectives in phonological studies (Hale & Reiss 1998: 658):

(1) The disparity between children's perception and production is attributed to:
   a. *grammatical* differences between children and adults, or
   b. *performance* effects such as motor processing.

The view in (1a) holds that child phonology is regarded as being fundamentally *distinct* from mature phonology in terms of the types of principles involved. This is widely accepted in the literature in various forms[1] and is supported by the observation that a child's production is regularity-governed (Smith 1973, Fee 1995).

Under the view in (1b), on the other hand, child phonology is subject to the *same* computational aspects as adult phonology and systematically observed regularities are properties of performance systems. This implies that the disparity stems from the difference in the degree of maturation in the performance systems of adults and children. This hypothesis is supported by Hale & Reiss (1998), who present empirical evidence and learnability considerations within an Optimality-theoretic framework.

I assume that, to a certain extent, both views in (1) are involved in the mismatch between perception and production in children. Limiting the argument to melodic contrasts at least, however, this paper in principle favours the view in (1b), that the difference between targets and the corresponding phonetic forms produced by a child must be the result of immature transmission facilities which cannot decode certain structures present in lexical representations. In other words, it is suggested that the mismatch between the child's production and the target form is attributed to the relation between the grammatical structures in lexical representations and immature performance facilities. This hypothesis underlies the view that performance facilities are innately equipped with some device for decoding phonological representations: these representations consist purely of abstract entities and provide no performance instructions or information for physical action and reception. This

---

[1] The literature includes the claims that the child manipulates two lexicons (Menn & Matthei 1992) and that it has drastically underspecified lexical representations (Ingram 1974).

view further implies that infants have access to the full set of innately given universal melodic categories, and that their lexical representations are more or less identical to those of adults.

Within this theoretical context, the following sections will analyse the acquisition of laryngeal-source contrasts in Japanese infants.

3. Laryngeal-source contrasts in Japanese

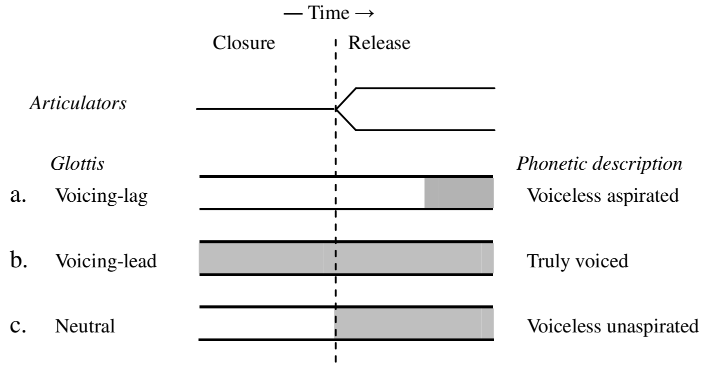
Before discussing the acquisition data, this section presents a cross-linguistic overview of laryngeal-source contrasts and shows how these operate in Japanese. In Japanese obstruents we find a two-way laryngeal-source contrast, which will be demonstrated here using word-initial plosives. This language is typical in allowing the widest possible distribution of consonant contrasts in this position; conversely, sounds in word-initial position rarely undergo deletion or neutralisation during derivation (Nasukawa & Oishi 2001: 133-134, cf. Beckman 1997: Ch1).

Laryngeal-source contrasts are often depicted in terms of voice onset time (VOT), which refers to the timing relation between the release of stop closure and the beginning of vocal-fold vibration in the following vowel (Lisker & Abramson 1964, Abramson & Lisker 1970, cf. Harris 1994, 1998).

(2) VOT

Plosive plus vowel sequence.

Glottal trace: open lines = vocal folds abducted; shading = vocal folds adducted.



Limiting the argument to languages exhibiting a two-way laryngeal-source contrast, studies of VOT reveal that there are two distinct ways of creating this type of contrast. One of the ways is found in languages such as English and Swedish, and bases the contrast on the distinction between long voicing lag (variously referred to as positive VOT, voiceless aspirated or fortis: 2a) and zero/short voicing lag (alternatively called neutral VOT, voiceless unaspirated, neutral or lenis: 2c). The other way is taken up in languages such as Spanish and Polish, and shows a contrast between long voicing lead (also known as negative VOT, voiced unaspirated or truly voiced: 2b) and zero/short voicing lag (2c). In both cases, neutral VOT is always involved: no language forms two-way laryngeal-source contrast without incorporating the neutral VOT.

(3) Language systems with a two-way laryngeal-source contrast

Type	E.g.	Neutral	Voicing-lead	Voicing-lag
I	English	✓	▨	✓
II	Spanish	✓	✓	▨

Japanese, the language under investigation here, belongs to the same group as Spanish and Polish (Type II in 3), exhibiting a contrast between long voicing lead and zero/short voicing lag. This classification of contrast types is supported by the results of perceptual tests for laryngeal-source discrimination (Shimizu 1977, 1996), and by the observation that, like Spanish and the other languages of the Type II group, aspiration (which is one characteristic of the two-way laryngeal-source contrast in Type II languages) is untraceable in the favoured contexts for VOT measurement in Japanese.

4. The acquisition of laryngeal-source contrasts in Japanese infants

4.1. The developmental disparities among performance systems in infants

In the literature on phonological development, a disparity is often found between the two domains of production and perception. A typical example would show how an infant can discriminate between two distinct words such as ‘mouth’ and ‘mouse’, even though these are neutralised (as *maʊt*) in the child’s own speech production.

As I shall argue below, however, the disparity is in fact observed across three distinct domains of performance – namely, the two domains just mentioned (production and perception of sound contrasts in a speaker’s utterance) and additionally the perception of sound contrasts in the infant’s own speech (self-monitoring of pronunciation). Evidence for the existence of self-monitoring comes from the fact that speakers need to apply this ability in order to utter sentences; conversely, the loss of this ability causes the speaker to produce incorrect pronunciations. Also, self-monitoring is crucial for delivering pragmatic effects before uttering an intentional sentence.

The following sections discuss the developmental paths of laryngeal-source contrasts in Japanese infants within these three domains.

4.2. Production

At the earliest stage of language acquisition, infants universally collapse all laryngeal-source contrasts into the region of zero/short voicing lag (Jakobson 1968: 14, *et passim*). This is supported, for example, by the results of studies on the acquisition of laryngeal-source based on experimental data from English and Spanish (Macken 1980). In the case of languages such as English, which create a contrast between zero/short voicing lag and long voicing lag, the latter property is acquired after the stage during which laryngeal-source contrasts are merged at the region of zero/short voicing lag. On the other hand, in languages such as Spanish, which exploit the contrast between zero/short voicing lag and long voicing lead, the convergence of laryngeal-source contrasts takes place around the region of zero/short voicing lag at the first stage of language acquisition and the contrast emerges at a later stage.

Japanese, which belongs to Type II in (3), also exhibits the same acquisition processes as other Type II languages do. Yet in fact, the process turns out to be less straightforward than is widely acknowledged. As mentioned in §4.1, there are not two, but three distinct stages in the acquisition of laryngeal-source contrasts:

(4) The acquisition of laryngeal-source contrasts in Japanese

a.	STAGE I (Kozumo: ~ 1:7, Nishiki: ~ 1:7):			
	Adult forms	e.g.	Infant forms	e.g.
	Short voicing lag *p*	*papa*	Short voicing lag *p*	***pa:pa***
	Long voicing lead *b*	*baba*	Short voicing lag *p*	***pa:pa***
b.	STAGE II (Kozumo: 1:8 ~ 2:0, Nishiki: 1:8 ~ 1:9):			
	Adult forms	e.g.	Infant forms	e.g.
	Short voicing lag *p*	*papa*	Short voicing lag *p*	***pa:pa*** ○ ○
	Long voicing lead *b*	*baba*	Short voicing lag *p*	***pa:pa***
c.	STAGE III (Kozumo: 2:0 ~, Nishiki: 1:10 ~):			
	Adult forms	e.g.	Infant forms	e.g.
	Short voicing lag *p*	*papa*	Short voicing lag *p*	***pa:pa***
	Long voicing lead *b*	*baba*	Long voicing lead *b*	***ba:ba***

The table in (4)[2] is based on a study of laryngeal-source distinctions in speech production, involving a longitudinal 'diary' investigation of two Japanese male infants: Kozumo (born on 16 June 2000) and Nishiki (born on 13 March 2002). Since their birth, they have been observed by the author (their father) who has kept transcribed records of their raw speech and made occasional audio recordings when required.

---

[2] My accumulated data also include another fourteen pairs of words containing two identical non-labial plosives: e.g. *tɕi:tɕi* 'wee-wee' - *dʑi:dʑi* 'grandpa' and *ka:ka* 'caw (of crows)' - *ga:ga* 'quack (of ducks)'.

With respect to the stages in (4), the first stage (4a) is the same as the literature often informs: the two-way laryngeal-source contrast present in adult forms converges in the short-lag region in infant speech. The forms consisting of the second stage (4b) have not been noted in the existing literature. In this stage, besides the convergence of laryngeal-source contrasts in the short voicing region, vowels followed by plosives with short voicing lag in adult forms are partially devoiced in infant production.[3] On the other hand, this kind of phenomenon is unattested in vowels after plosives with long voicing lead in infant forms. That is, the lexical contrast is observed not on the plosives themselves, but on the following vowels. In the final stage (4c), the adult-like laryngeal-source contrasts emerge.

The above aspects in the three stages are common between the two infants. There is, however, at least one feature which was observed in Nishiki but not in Kozumo: between Stages II and III, Nishiki produced the form *ha:pa* where there is no phonetic indication of devoicing on the vowels and which is distinct from *ba:ba*.

4.2. Perception

The perception of laryngeal-source contrasts is, as discussed in §4.1, classified into two types: the capacity for identifying sound contrasts in speakers' utterances (Type A) and the monitoring of self-pronunciation (Type B). It is widely acknowledged that,[4] unlike the production phase, the ability to identify the two-way laryngeal-source distinction is already established at a very early stage of acquisition (cf. Jusczyk 1998). On the other hand, the infant's ability to discriminate his or her own pronunciation, the Type B ability, is not established at Stage I. This self-monitoring ability is developed in Stage II, where infants, while maintaining the ability A, begin to distinguish their own speech production. At Stage III, the Type B ability finally matures into something matching its adult equivalent.

---

[3] The devoiced vowels following the child's plosive form corresponding to its adult form with short voicing lag may be regarded as whispered sounds, something which is never found in the vowels following the child's plosive form corresponding to its adult form with long voicing lead.

[4] The perceptual ability of voicing distinctions is investigated simply by asking the infants to answer yes-no questions.

(5) The diachronic development of laryngeal-source distinctions in Japanese infants

	PRODUCTION	PERCEPTION	
		TYPE A	TYPE B (Self monitoring)
STAGE I	**Not established:** Convergence at short-lag region	**Established**	**Not established:** Unable to self-monitor
STAGE II	**Not established:** Convergence at short-lag region. Vowel devoicing.	**Established**	**Established**
STAGE III	**Established**	**Established**	**Established**

Below I will investigate the differences in the diachronic development of laryngeal-source distinctions in (5), and relate these findings to the phonological representations of element-based feature theory.

5. Element-based melodic theory

Element Theory (ET: Kaye, Lowenstamm & Vergnaud 1985; Harris & Lindsey 1995, 2000 and references therein) is characterized by the use of a set of features called elements. These elements differ from those features employed in orthodox distinctive feature theories such as SPE (Chomsky & Halle 1968) by being privative in the sense of Trubetzkoy (1939); this means that they may be interpreted separately without needing to be combined with other elements. The result is the elimination of redundancy fill-in rules and predictable feature values, leading to an entirely monostratal approach to phonology (Harris 2004). The following list, which is based on Harris (1994, 2004) and Harris & Lindsey (1995), shows each element with its respective acoustic and articulatory interpretation.

(6) Elements

	ELEMENTS	PATTERN	ACOUSTIC INTERPRETATION	ARTICULATORY INTERPRETATION
a.	Resonance/Place			
	[A]	Mass	Central spectral energy mass (Convergence of F1 and F2)	Maximal expansion of oral tube; maximal constriction of pharyngeal tube
	[I]	Dip	Low F1 coupled with high spectral peak (Convergence of F2 and F3)	Maximal constriction of oral tube; maximal expansion of pharyngeal tube
	[U]	Rump	Low spectral peak (Convergence of F1 and F2)	Trade-off between expansion of oral and pharyngeal tubes
	[R]	Rise	High spectral peak	Articulation with the tip or the blade of the tongue (coronality)
	@	Neutral	No salient spectral peak	Neutral expansion of oral tube; neutral constriction of pharyngeal tube (centrality and velarity)
b.	Manner/Laryngeal			
	[?]	Edge	Abrupt and sustained drop in overall amplitude	Occlusion in oral cavity
	[h]	Noise	Aperiodic energy	Narrowed stricture producing turbulent airflow
	[N]	Murmur	Low broad resonance peak	Lowering of the velum
	[H]	high source	F0 up	Spread vocal folds

The inventory in (6) contains no independent element corresponding to the feature [voice] – in the interests of reducing the overall size of the element set, an autonomous element such as [L] has been eliminated from the inventory.[5]

(7) Element reduction:
Backley (1993) for eliminating [R]
Harris & Lindsey (1995) for eliminating [ATR]
Nasukawa (1995, 1998, 2000) for eliminating [L] as a true voicing category
Yoshida (1995) for eliminating [L] as a low tone category
Cabrera-Abreu (2000) for eliminating [L] as a low tone category

According to the theory, long voicing lead (true voicing) is a phonetic manifestation of the preponderance of the murmur element [N] over the other element(s) with which it is combined; furthermore, a necessary condition for

---

[5] In Ploch (1999) and Kaye (2001), [L] remains in the element inventory instead of [N].

specifying a headed [N] is the presence of the noise element [h]. This was first formally investigated by Nasukawa (1995, 1998, 2000), while ample supporting evidence and similar arguments can be found in Ploch (1999), who employs [L] in place of [N].[6]

(8) Element representation Phonetic manifestation
a. non-headed [N] => nasality is phonetically discernible
b. headed [$\underline{\text{N}}$] + [h] => long voicing lead is phonetically discernible
(underlining implies headedness)

When [N] is not headed in a given melodic expression, it is phonetically interpreted as nasality. This notion of headedness is fundamental to the element-based approach to melodic organization (for a detailed discussion, see Harris 1994 and Kaye 1995) and is different from that in prosody and the other components of linguistic competence. In prosody and higher levels, the notion of headedness is called upon for describing dependency relations between two categories, while the same idea in ET translates into the preponderance over not only a single category, but possibly over two or more categories which comprise a given melodic expression. In terms of acoustic realization, melodic headedness in ET functions to enhance the acoustic properties of a particular element over those of other element(s).

In order to clarify the difference between headedness in melody and in other domains, Backley & Takahashi (1998) incorporates a melodic complement tier ([comp]), which may be licensed by a single element within a given expression and has the effect of enhancing the acoustic image of its licensing element over the whole expression. Another difference between [comp] and headedness in standard ET is their contribution to melodic complexity: the latter makes no contribution to melodic complexity while the former is counted on a par with other elements in terms of complexity. This is in spite of its lack of any stable phonetic signature and its auxiliary role of merely enhancing the acoustic image of its licensing element. Below, I will follow standard ET practice by representing headedness using underlining, rather than [comp]; however, I will treat it like [comp] in the sense that it contributes to melodic complexity in the way other elements do.

Now using the elements described so far, plosives – which typically exhibit VOT-based contrasts – are represented as follows:

---

[6] I have opted for [N] rather than [L] to represent the relation between nasality and long-lead voicing, since the 'bare' element without its headedness contributes nasality. Another reason to use [N] rather than [L] for presenting the correlation between nasal and voice is attributed to a recent study of tone and intonation, in which low tone (corresponding to [L] in ET in nuclear position) is the phonetic manifestation not of an independent melodic category but of a prosodic boundary not associated to high tone (corresponding to [H] in ET in nuclear position) (Cabrera-Abreu 2000).

(9) a. Bilabial plosives
i. *p* ii. *b*
[U, ?, h] [U, ?, h, N]
b. Coronal plosives
i. *t* ii. *d*
[R, ?, h] [R, ?, h, N]
c. Velar plosives
i. *k* ii. *g*
[@, ?, h] [@, ?, h, N]

As (9) shows, plosives usually contain [?] and [h]: in fact, the cooccurrence of [?] and [h] in a single melodic expression is necessary to form a plosive segment. In terms of articulatory execution, the resonance elements [U], [R] and @ (= the non-specification of any resonance elements) contribute labiality, coronality and velarity respectively in non-nuclear positions. Concerning laryngeal-source contrasts, members of the neutral plosive series (9ai, 9bi, 9ci) are thus composed of only three elements – [?], [h] and one of the resonance elements (and in the case of the velar plosive, no specified resonance element); there is no element specification for laryngeal-source contrasts. On the other hand, members of the long-voicing-lead series (9aii, 9bii, 9cii) are represented not only by the three elements but also by an additional headed [N], which contributes long voicing lead. That is, the two-way laryngeal-source contrast in Japanese is represented by the contrast between the absence of any voicing element and the specification of a headed [N].

The additional category [N] straightforwardly captures those assimilatory processes involving laryngeal activities in Japanese – and in particular, those processes triggered by true voicing (the property of long voicing lead). For example, in the process of sequential voicing known as Rendaku, a headed [N], which is the compounding conjunction morpheme, links to the initial obstruent (containing [h]) of the second member of a compound unless the process is blocked by Lyman's Law (Nasukawa 1998, 2000; cf. Itô & Mester 1986). In the case of postnasal voicing, [N] in the nasal of an NC sequence manifests itself in the following obstruent containing [h] and receives headed status in accordance with the complexity condition (Nasukawa 2000, 2003; cf. Harris 1990, 1994). On the other hand, the non-specification of voicing elements in neutral members means that phonation properties are predicted to be phonologically inert. According to Nasukawa (to appear), vowel devoicing in Japanese, which is traditionally considered to be brought about by 'voicelessness', is not a process triggered by any voicing element (such as [H] for aspiration employed in languages like English), but by the noise element [h].

6. Complexity and interpretability

As noted in §2.2, I take the position that the mismatch between a child's production/Type B perception and its target is attributed to the relation between structures in lexical representation, which are governed by grammar, and immature performance facilities. According to this view, the representation of Japanese plosives in infants is identical to that found in adult forms. So in (9), for instance, *p* and *b* consist of [U, ?, h] and [U, ?, h, N] respectively.

The mismatch between a phonological form and its phonetic interpretation, as discussed in §2.2, may be attributed to the under-developed reception and transmission facilities, which interpret phonological representations and map them onto phonetic entities. The following table shows how phonological representations are interpreted by the reception and transmission facilities at each stage.

(10) Representation of laryngeal-source distinctions by Japanese infants

		Production	Perception	
			Type A	Type B
Stage I	*p*	[U,?, h]	[U,?, h]	[U,?, h]
	b	[U, ?, h, <N>]	[U, ?, h, N]	[U, ?, h, <N>]
Stage II	*p*	[U,?, **h**]	[U,?, h]	[U, ?, h]
	b	[U, ?, h, <N>]	[U, ?, h, N]	[U, ?, h, N]
Stage III	*p*	[U,?, h]	[U,?, h]	[U, ?, h]
	b	[U, ?, h, N]	[U, ?, h, N]	[U, ?, h, N]

As for the interpretability of the reception facility, Type A perception is identical to that in adults at all stages of development. With respect to Type B perception, on the other hand, the headed murmur element [N] is not perceived (< > implies 'not interpreted') at Stage I, whereas in the following stage, it is successfully perceived along with all other elements.

In the case of the development of the transmission facility, like the Type B perception ability, [N] is not interpreted; therefore, its corresponding phonetic signature is not present in infant speech at Stage I. After Stage II, the developed perception ability Type B prevents any convergence of the laryngeal-source contrasts into the short-lag region in infant speech. However, since the production ability has not reached a sufficiently steady state for delivering laryngeal-source contrasts, interpreting [N] instead, [h] in the representation of the plosive with short voicing lag is enhanced and is phonetically realized in the following nuclear position. The point at which the production ability reaches its steady state is Stage III, where all other abilities function as in adult speech.

The above argument raises at least the following two questions:

(11) a. Why is [N], rather than the other elements, not interpreted by the reception and transmission facilities in the early stages?

b. When [N] cannot be interpreted by the transmission facility, why is [h] – rather than some other element, selected for delivering lexical contrasts at Stage II?

The answer to (11a) is related to the melodic organisation of plosives in Japanese. Compare the short voicing lag plosive with the long voicing lead cognate: an extra element [N] is specified in the latter representation. In addition, the category possesses the property of headedness which, in a recent version of Element Theory (Nasukawa 1999, 2000, to appear), is employed as a property which adds complexity to a given structure. Further, I claim that a melodically complex expression is difficult for the performance systems to interpret during the earliest stages of language acquisition. The same phenomenon is observed in nuclear systems: ATR-ness, which is represented by headedness on resonance elements (e.g. *i* consists of [I] in comparison with its lax counterpart *ɪ* consisting of [I]), is acquired at a later stage.

In response to the question in (11b), this is associated with the relation between [N] and [h]. As briefly stated in §5, in order to interpret a headed [N] structure, the noise element [h] must also be specified in the same expression. This may be explained by the fact that obstruents – which are characterized by the obligatory presence of the noise element [h] – provide the most natural context for supporting laryngeal-source contrasts. This relation between [h] and [N] is formalised as follows (cf. N-[COMP] LICENSING within the model of Element geometry in Nasukawa 2000: 103-104):

(12) **[N]-HEADEDNESS LICENSING**
A headed [N] must be licensed by [h].

I assume that this tight relation between [N] and [h] makes infants select [h] as a category for delivering lexical contrasts in their speech when [N] is not available. In this case, [h] in the representation of the plosive with short voicing lag (*p*), rather than in that with long voicing lead (*b*), becomes prominent in its expression. Its phonetic signature manifests itself in the following nuclear position and delivers the effect of devoicing the lexically-given vocalic quality at Stage II. This kind of devoicing commonly occurs in vowels positioned between 'voiceless' consonants in Japanese (Nasukawa to appear). An extreme case of [h]-enhancement is found in the form *ha:pa* produced by Nishiki between Stages II and III, where only the enhanced element [h] remains and all other elements are suppressed.

The assumption that [h] in the form of *p* is enhanced is attributed to the ET's structural restriction that only one element may be licensed as a head within any melodic expression. In the representation of *b*, headedness is assigned to [N], where [N] is the only element to have its acoustic signature phonetically enhanced in this way. The representation of *p*, on the other hand, has no element licensing the headedness property. In such a case, [h], which is associated to the property which unites all obstruents and which delivers laryngeal-source contrasts, is free to be enhanced and can be acoustically prominent.

7. Conclusion

To conclude, the acquisition of the sound properties of human language may be attributed to an interaction between the phonological component and the reception and transmission facilities in the language faculty. Particular aspects of phonological representation cannot be decoded by performance systems at an early stage of acquisition. Here, the notion of difficulty, which is a property of the transmission and reception facilities (not a property of phonological competence) is involved: it is difficult for immature transmission and reception facilities to interpret complex structures in a particular domain/level. In the case of the acquisition of laryngeal-source contrasts in Japanese infants, the self-monitoring ability of the reception facility is unable to interpret the headed murmur element [$\underline{N}$] at Stage I; additionally, the transmission facility is unable to interpret [$\underline{N}$], not only at stage I but also at Stage II. The reason why [$\underline{N}$] is not interpreted by those facilities during the early stages of acquisition may be explained by referring to the complexity of [$\underline{N}$] in melodic representation. The development of laryngeal-source contrasts in those facilities finally reaches Stage III, when infants are capable of interpreting all phonologically-specified categories.

In order to support the arguments presented here, further investigation is required into the acquisition of other sound properties (including other melodic contrasts and prosodic phenomena), not only in Japanese infants but also in a wider range of other child languages.

Acknowledgements

I am grateful to Marilyn Vihman, Joe Pater, René Kager, and Nancy Chongo Kula for their kind and valuable comments on this paper, and to the audience at Developmental Paths in Phonological Acquisition, Leiden University, 16-18 September 2004. This paper has benefited greatly from insightful comments and suggestions from Neil V. Smith and Phillip Backley. Research for this paper was partially funded by the Ministry of Education, Culture, Sports, Science and Technology of the Japanese government under grant 16720117.

Kuniya Nasukawa
Department of Letters, Tohoku Gakuin University,
Tsuchitoi 1-3-1, Aoba-ku, Sendai, 980-8511,
Japan.
e-mail: nasukawa@tscc.tohoku-gakuin.ac.jp

References

Abramson, A. S. & L. Lisker. (1970). Discriminability along the voicing continuum: cross-language tests. *Proceedings of the 6th International Congress of Phonetic Sciences*. Academia, Czechoslovak Academy of Sciences, Prague, pp. 569-573.

Backley, P. (1993). Coronal: the undesirable element. *UCL Working Papers in Linguistics* 5. pp. 301-323.

Backley, P. & T. Takahashi (1998). Element activation. E. Cyran (ed.), *Structure and interpretation: studies in phonology*. Folium, Lublin, pp. 13-40.

Beckman, J. N. (1997). *Positional faithfulness*. Doctoral dissertation, University of Massachusetts, Amherst.

Braine, M. D. S. (1976). Review of N. V. Smith, *The acquisition of phonology; a case study*. *Language* 52, pp. 489-498.

Cabrera-Abreu, M. (2000). *A phonological model for intonation without low tone*. Indiana University Linguistics Club Publication, Bloomington, Indiana.

Chomsky, N. & M. Halle (1968). *The sound pattern of English*. Harper and Row, New York.

Fee, E. J. (1995). Segments and syllables in early language acquisition. J. Archibald (ed.), *Phonological acquisition and phonological theory*, Lawrence Erlbaum Associates, Hillsdale, New Jersey, pp. 43-61.

Hale, M. & C. Reiss (1998). Formal and empirical arguments concerning phonological acquisition. *Linguistic Inquiry* 29, pp. 656-683.

Harris, J. (1990). Segmental complexity and phonological government. *Phonology* 7, pp. 255-301.

Harris, J. (1994). *English sound structure*. Blackwell, Oxford.

Harris, J. (1998). Phonological universals and phonological disorder. E. Visch-Brink & R. Bastiaanse (eds.), *Linguistic levels in aphasia: Proceedings of the RuG-SAN-VKL Conference on Aphasiology*. Singular Publishing Group, San Diego, CA, pp. 91-117.

Harris, J. (2004). Release the captive coda: the foot as a domain of phonetic interpretation. J. Local, R. Ogden & R. Temple (eds.). *Phonetic interpretation: Papers in Laboratory Phonology* 6. Oxford University Press, Oxford, pp. 103-129.

Harris, J. & G. Lindsey (1995). The elements of phonological representation. J. Durand & F. Katamba (eds.), *Frontiers of phonology: atoms, structures, derivations*. Longman, Harlow, Essex, pp. 34-79.

Harris, J. & G. Lindsey (2000). Vowel patterns in mind and sound. N. Burton-Roberts, P. Carr & G. Docherty (eds.), *Phonological knowledge: conceptual and empirical issues*. Oxford University Press, Oxford, pp. 185-205.

Ingram D. (1974). Phonological rules in young children. *Journal of Child Language* 1, pp. 49-64.

Itô, J. & R-A. Mester (1986). The phonology of voicing in Japanese. *Linguistic Inquiry* 17, pp. 49-73.

Jakobson, R. (1968). *Child language, aphasia and phonological universals*. Mouton, The Hague.

Jusczyk, P. W. (1997). *The discovery of spoken language*. The MIT Press, Cambridge, Massachusetts.

Kaye, J. D. (1995). Derivations and interfaces. J. Durand & F. Katamba (eds.), *Frontiers of phonology: atoms, structures, derivations*, Longman, Harlow, Essex, pp. 289-332.

Kaye, J. D. (2001). Working with Licensing Constraints. Ms, Guangdong University of Foreign Studies.

Kaye, J. D., J. Lowenstamm & J-R. Vergnaud (1985). The internal structure of phonological representations: a theory of charm and government. *Phonology Yearbook* 2, pp. 305-328.

Lisker, L. & A. S. Abramson (1964). A cross-language study of voicing in initial stops: acoustical measurments. *Word* 20, pp. 384-422.

Macken, M. A. (1980). Aspects of the acquisition of stop systems: a cross-linguistic perspective. G. H. Yeni-Komshian, J. F. Kavanagh & C. A. Ferguson (eds.), *Child phonology volume 1: production*. Academic Press, New York, pp. 143-168.

Mehler, J., E. Dupoux, T. Nazzi & G. Dehaene-Lambertz (1996). Coping with linguistic diversity: the infant's viewpoint. J. Morgan & K. Demuth (eds.), *Signal to syntax: bootstrapping from speech to grammar in early acquisition*. Lawrence Erlbaum Associates, Hillsdale, New Jersey, pp. 101-116.

Menn, L. & E. Matthei (1992). The 'two-lexicon' account of child phonology: looking back, looking ahead. C. Ferguson, L. Menn & C. Stoel-Gammon (eds.), *Phonological development: models, research, implications*. York Press, Timonium, MD, pp. 211-247.

Nasukawa, K. (1995). Melodic structure and no constraint-ranking in Japanese verbal inflexion. Paper presented at the Autumn Meeting of the Linguistic Association of Great Britain. University of Essex, September.

Nasukawa, K. (1998). An integrated approach to nasality and voicing. E. Cyran (ed.), *Structure and interpretation: studies in phonology*. Folium, Lublin, pp. 205-225.

Nasukawa, K. (1999). Prenasalisation and melodic complexity. *UCL Working Papers in Linguistics* 11, pp. 207-224.

Nasukawa, K. (2000). *A unified approach to nasality and voicing*. Doctoral dissertation, University College London, University of London. [Published in 2005 by Mouton de Gruyter, Berlin/New York.]

Nasukawa, K. (2003). Direction of assimilation reflects prosody-melody interaction. *Proceedings of the UCL Alumni Reunion Conference*. University College London, University of London, pp. 29-33.

Nasukawa, K. (to appear). The representation of laryngeal-source contrasts in Japanese. J. van de Weijer, K. Nanjo & T. Nishihara (eds.), *Voicing in Japanese*. Mouton de Gruyter, Berlin/New York.

Nasukawa, K. & M. Oishi (2001). Inaccessibility of the domain-initial nucleus in high-pitch agreement. M. C. Cuervo, D. Harbour, K. Hiraiwa & S. Ishihara (eds.), *Formal Approaches to Japanese Linguistics* 3. *MIT Working Papers in Linguistics* 41, Massachusetts Institute of Technology, pp. 127-144.

Ploch, S. (1999). *Nasals on my mind: the phonetic and the cognitive approach to the phonology of nasality*. Doctoral dissertation, School of Oriental and African Studies, University of London.

Shimizu, K. (1977). Voicing features in the perception and production of stop consonants by Japanese speakers. *Studio Phonologica* 11, pp. 25-34.

Shimizu, K. (1996). *A cross-language study of voicing contrasts of stop consonants in Asian languages*. Seibido Publishing Co., Tokyo.

Smith, N. V. (1973). *The acquisition of phonology: a case study*. Cambridge University Press, Cambridge.

Smith, N. V. (2003). Are gucks mentally represented? *Glot International* 7, pp. 164-166.

Smolensky, P. (1996). On the comprehension/production dilemma in child language. *Linguistic Inquiry* 27, pp. 720-731.

Trubetzkoy, N. S. (1939). *Grundzüge der phonologie. Travaux du cercle linguistique de Prague* 7.

Yoshida, Y. (1995). *On pitch accent phenomena in standard Japanese*. Doctoral dissertation, School of Oriental and African Studies, University of London. [Published in 1999 by Holland Academic Graphics. The Hague.]

Title: Melodic complexity and disparities between child’s performance and the target

Author: Kuniya Nasukawa
Department of Letters, Tohoku Gakuin University,
Tsuchitoi 1-3-1, Aoba-ku, Sendai, 980-8511,
Japan.

e-mail: nasukawa@tscc.tohoku-gakuin.ac.jp