Gestural Hesitation Reveals Children's Competence on Multimodal Communication: Emergence of Disguised Adaptor

Kazuki Sekine (kazuki.sekine@mpi.nl)
Radboud University/ Max Planck Institute for Psycholinguistics,
Wundtlaan 1, 6526 XD Nijmegen, the Netherlands

Abstract

Speakers sometimes modify their gestures during the process of production into adaptors such as hair touching or eye scratching. Such *disguised* adaptors are evidence that the speaker can monitor their gestures. In this study, we investigated when and how disguised adaptors are first produced by children. Sixty elementary school children participated in this study. There were ten from each school year (from 7 to 12 years of age). They were instructed to remember a cartoon and retell its story to their parents. The results showed that children did not produce disguised adaptors until the age of 8. The disguised adaptors accompany fluent speech until the children are 10 years old and accompany dysfluent speech until they reach 11 or 12 years of age. These results suggest that children start to monitor their gestures when they are 9 or 10 years old. Cultural influences and cognitive changes were considered as factors to influence emergence of disguised adaptors.

Keywords: co-speech gestures; disguised adaptors; elementary school children; speech dysfluency.

Introduction

Researchers have examined the development of gestures in children in terms of when the frequency or repertoire increases and how the relationship between a gesture and speech changes with age. The present study focused on spontaneous suppression of gesture production during speech. Analyzing when and how children try to not produce gestures would provide insight into when children become aware that their gestures are socially communicative.

The present study focuses on co-speech gestures (hereafter simply referred to as 'gestures') that spontaneously co-occur during speech and that have no standard of well-formedness, unlike sign language, but are created idiosyncratically on the fly (McNeill, 1992). A gesture typically has three phases: *preparation*, *stroke*, and *retraction*. In the preparation phase, the hand moves from a position of rest. The stroke phase is the central part of a gesture that conveys substantial information. The relevant meaning represented by this phase is usually expressed in the concurrent speech. Sometimes, a hold phase, where the hand is held in mid-air at the same position, occurs before and/or after the stroke phase.

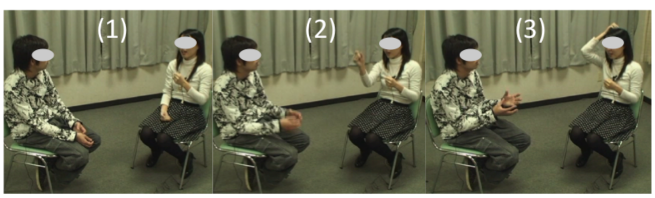
As children prefer to use speech as a means of communication, the frequencies of gestures that appear around the first word and are used alone without speech, such as deictic gestures and symbolic gestures, decreases (Volterra & Iverson, 1995). In their place, gestures that co-occur with meaningful words, called "*co-speech gestures*", appear around the period. Goldin-Meadow & Butcher (2003) observed that gestures begin synchronizing with speech both semantically and temporally in the transitional phase to the two-word speech period, at about 18 months of age. Given that gestures are not often used solely but are co-produced with speech, gestures and speech seem to form an integrated system during this period (Goldin-Meadow & Butcher, 2003).

Previous research has shown that speech and co-speech gesture develop hand-in-hand even after two-word period. Mayberry & Nicoladis (2000) observed that longitudinally bilingual children between 2- and 3-and-half years old produce more gestures when they speak either language that allows them produce a longer utterance, as measured by the mean number of morphemes (Mayberry & Nicoladis, 2000). They concluded that gestures are closely related to morphosyntax level. This conclusion is indirectly supported by Fujii's (1999) study showing that the frequency of gestures does not correlate with vocabulary in the preschool period. McNeill (1992) observed that by the end of the preschool period, the frequency of gestures ascends to near adult levels. Once children start having formal education, they gradually develop the ability to create a coherent narrative by using language devices such as anaphora expressions, substitutions, ellipses and connectives (Wigglesworth, 1990). Research has shown that as children acquire spoken referential expressions for making coherent discourse, they also use gestures to mark introduced or maintained referents in the narrative (McNeill, 19992), and the number of gestures consistently increases during the elementary school period (Colletta et al., 2014; Sekine & Furuyama, 2010). Thus, these previous studies indicate that gesture and speech develop hand-in-hand across the development. However, it is not clear if children start monitoring their gestures or become aware of gestures as information resources that their listeners can make use of.

Studies on self-repair in speech have asserted that to correct one's speech, the speaker has to monitor his or her speech process continuously, and thus, correction of speech errors is considered to reflect the speaker's ability to monitor speech (Karmiloff-Smith, 1986). Based on this assumption, it seems that determining when children start to correct their gestures would provide insight into understanding when they start monitoring their gestures. However, unlike speech in which the message is delivered by aligning linearly linguistic elements that exist independently, a gesture conveys a meaning globally at once, and any decomposition into elements is dependent on the whole (McNeill, 1992). Because of the linearity in speech, it is easier to understand where speech errors occurred and how the speaker corrected them than errors in gestures. In contrast, gestures are mostly continuous and some of their parameters change simultaneously. This makes it difficult to determine whether and where a speaker has corrected a gesture. Considering these differences in semiotic characteristics, this study focused on a specific type of gestural correction; i.e., *disguised adaptors.*

Adaptors are movements that help persons to satisfy personal needs, manage emotions, and adapt to their environment such as touching one's hair or adjusting one's glasses (Ekman & Friesen, 1969). At times it is observed that a speaker stops making a gesture in the middle of production and switches it to an adaptor to hide the gesture. Such movements may be able to say socially preferable in the situation. In this study, this kind of behavior is termed a *disguised adaptor*, which is defined as a gesture that is altered into an adaptor before or during the stroke phase of the gesture.

It can be observed in daily conversation that when a speaker is asked a question by the listener in the middle of her gesture and speech production, she stops producing the gesture, and puts her hand on her head or eye to scratch as if it feels itchy, like shown in Figure 1. Figure 1 indicates a scene in which the speaker on the right was retelling an episode of cartoon that she had watched to the listener on the left ((1) in Figure 1). When she was describing the figure of the drainpipe with a gesture, the listener started asking her whether one character was in the birdcage (2). At that moment, she stopped her gesture in the middle, and put her right hand on her head and scratched it until the listener finished the question (3).



Speaker (on the right side):
[soko made (1)**ikitakutte**, sono mado no (2)**tokoro nikou-**
'(he) (1)wants to go there, and (2)at the place where the window is , like this-'
Listener (on the left side):
[torikago (3)**no naka ni haitteiruno**?]
'is (he) (3)inside the cage?'

Figure 1: Halt of a gesture by listener's question.

Here, and in subsequent examples, the square brackets represent the start and end points of the motion of the speaker's hands, boldface marks the stroke phrase of the gesture phrase, underline indicates a motionless hold phase, and double underline represents the duration in which a disguised adaptor, such as touching the body or clothes. '%' indicates a smacking sound, '*' represents self interruption, and ':' in speech indicates an elongated phonation. The numbers used in figure, correspond to the numbers in the transcription, which in turn indicate the places where gestures occurred. In the transcription, the first chunk is the original Japanese speech and the second chunk is the English translation.

Similar behaviors to the one in Figure 1 can be observed in other situations, for example when one raises his hand to catch a taxi, but misses it, or in which one is waving to her friend, and then quickly becomes aware of mistaken identity. In these situations, they often switch the hand movements to self-contact behaviors such as scratching head or eyes as if it is meant to be. Interestingly, literature on Tourette's syndrome, a chronic neurological disorder characterized by multiple involuntary movements and uncontrollable vocalizations called tics, has documented the correction of tic movements made by patients. Sottofattori and Nicolai (2007) observed that the patients can modify tics or odd movements into other movements like gestures in a natural conversation. For example, when a tic affects the right or left arm, the patient tries to move the forefinger straight with the tic as if it were a deictic gesture. All these cases suggest that when we are aware of mistake of hand movement or when we are interrupted in executing our hand movements, we often change the hand movement to a more socially acceptable movement such as scratching a body part or producing gestures.

In the light of development of gestures in children, it is important to examine whether typically developing children also are able to modify body movement into a more socially acceptable one. Because if children perform this kind of correction, it implies that they can monitor their gestures. If they already know that gestures can be seen as information resource by their listeners and they can detect errors in their gestures or speech, they may try to modify their gestures into more socially acceptable movement, such as scratching the neck.

Suppose that gesture and speech interacts in the production process (McNeill, 1992), and then gestural correction may affect or be affected by speech errors or dysfluencies. Thus, the relationship between disguised adaptors and speech errors should also be investigated. Karmiloff-Smith (1986) found that the percentage of speech repairs denoting sensitivity to the linguistic system, such as the determiner functions of articles, adjectives and possessives, increased during the elementary school period. Given these findings, it is predicated that the number of disguised adaptors also increases during this period as children acquire an ability to monitor their expressions. In addition, if disguised adaptors are due to difficulties in retrieving words, they would co-occur with filled pauses, unfilled pauses, or speech errors, rather than with intact linguistic elements, because previous studies have shown that adaptors tend to occur with speech dysfluencies while the speaker is retrieving words (Fujii, 1997; Pine, Bird, & Kirk, 2007). However, if a disguised adaptor does not co-occur with speech dysfluencies, it might occur influenced by other factors such as cultural or cognitive factors. For instance, cognitive capacity or cultural pressure to produce or inhibit gestures may control the emergence of disguised adaptors. Under this hypothesis, this study examined when disguised adaptors emerge during the elementary school period and how this emergence is related to speech.

Method

Participants

Sixty elementary school children and their parents participated in this study. There were ten children from each grade, 1st to 6th grade. Half were boys. In this study, each grade is referred to by their mean age, 7- to 12-years-olds (7-year-olds, M=7;0, Range 6;9-7;4. 8-year-olds, M=7;11, range 7;9-8;10, 9-year-olds. M=9;4, R= 9;0-9;6, 10-year-olds, M=10;0, R=9;7-10;5, 11-year-olds, M=11;4, R= 10;9-11;11, 12-year-olds, M=12;0, R=11;5-12;10). All the participants were native monolingual Japanese speakers from middle-class families, and the children attended public or private elementary schools in Tokyo, Japan.

Material and Apparatus

Each child watched a seven-minute animated color cartoon of the Tweety and Sylvester series, titled ‘Canary Row’ (Warner Brothers, Inc.). This cartoon was displayed on a 14-inch color computer monitor (Panasonic CF-F8). A mini-DV camcorder (Sony HDR-HC9) was used to record the children’s gesture and speech.

Procedure

The experiments were conducted in a quiet room in the participant’s home or a local community center. Each child was instructed to remember the cartoon story shown on the computer monitor and retell it to his or her parent as a listener in as much detail as possible. The parent was allowed to respond to the child by nodding their head or by using back channels during the child’s narration. The whole session was recorded using the mini-DV camcorder on a tripod.

Coding

Speech data. All narratives were verbatim transcribed. From the transcriptions, the *mean number of clauses* was then calculated. A clause was loosely defined as a combination of a noun phrase and a verb phrase). The *mean number of unfilled pauses*, which was defined as periods of silence longer than 200 msec. (Beattie, 1983), *filled pauses*, such as ‘unttoo’ (umm) or ‘eeetto’ (ehhh), and *speech errors*, including repetitions, replacements or false starts, were measured in order to ascertain the relationship between speech fluency and production of disguised adaptors.

Gesture data. First, co-speech gestures were identified. Hand movements were classified as gestures only when they had an identifiable beginning and a clear end and were synchronized with speech. After identifying which movements were gestures, the total number of gestures was counted. Next, *disguised adaptors* were identified. We coded a hand movement as a disguised adaptor if two criteria were met; 1) A gesture was altered into adaptor before or during the stroke phase without any pause, and 2) the direction of the gesture suddenly changed when it is altered into adaptor. To analyze the temporal relationship between disguised adaptors and speech fluency, their combinations were categorized into the following six types: DA (disguised adaptor), FP (filled, pause), IS (intact speech that was completely articulated, such as Noun or Verb phrases), PR (the preparation phase of a gesture), SE (speech error), and UP (unfilled pause). The orders were as follows:

1. PR → IS+DA
After the preparation phase started, a disguised adaptor occurred with intact speech.
2. PR → FP+DA
After the preparation phase started, a disguised adaptor occurred with a filled pause.
3. PR → UP+DA
After the preparation phase started, a disguised adaptor occurred with an unfilled pause.
4. PR → FS+DA
After the preparation phase started, a disguised adaptor occurred with a filled pause.
5. PR → FS (FS) →IS+DA
After the preparation phase started, a false start or a number of false starts occurred. Then, a disguised adaptor occurred with intact speech.
6. PR→ FP and/or FS → FP+DA
After the preparation phase started, a false start and/or filled pause occurred. Then, a disguised adaptor occurred with a filled pause.

Results

Number of gestures, clauses, pauses, and speech errors
To calculate the frequency of gestures and the proportion of unfilled pauses for each age, the total number of gestures and the total amount of time spent for unfilled pauses were divided by the total speaking time. To calculate the frequency of filled pauses and speech errors, the total number of filled pauses and speech errors were each divided by the mean number of clauses (Table 1). After performing an angular transformation on the proportion of unfilled pauses in the speaking time, an ANOVA was conducted on each index (Table 1). A main effect of age group was found for the gesture frequency, F (5, 54) = 3.86, p= .005, the proportion of unfilled pauses in the speaking time, F(5, 54) = 2.44 p= .046, and the proportion of speech errors in clauses, F(5, 54) = 2.49, p= .042. A post hoc comparison (Tukey, $p < .05$) showed that 12-year-olds produced gestures more frequently than 7-, 9- and 10-year-olds did and that the proportion of unfilled pauses during the speaking time was significantly greater for 7-year-olds (47%) than for 12-year-olds (33%). There was no significant age-group difference in the total number of clauses or frequency of speech errors. These results indicate that the proportion of unfilled pauses gradually decreases during the elementary school period, whereas the frequencies of

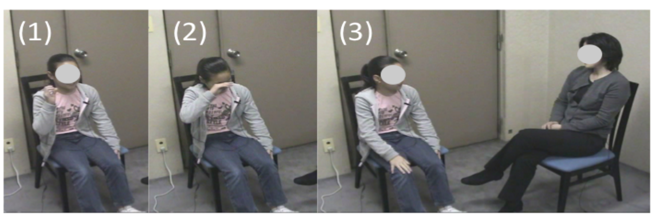
gestures and speech errors increase in the late elementary school period.

Table 1: Number of Clauses, Gestures, Pauses, and Speech Errors.

Age	7	8	9	10	11	12
-Frequency of gestures per second	0.12 (0.06)	0.14 (0.09)	0.10 (0.09)	0.09 (0.06)	0.19 (0.13)	0.27 (0.17)
-Total number of clauses	62.4 (29.93)	69 (20.42)	77.8 (24.68)	77.3 (15.58)	91.9 (20.90)	78.6 (19.47)
-Proportion of unfilled pauses in speaking time	0.47 (0.19)	0.40 (0.11)	0.38 (0.08)	0.34 (0.06)	0.36 (0.08)	0.33 (0.08)
-Filled pauses per clause	0.38 (0.25)	0.29 (0.17)	0.26 (0.07)	0.39 (0.31)	0.31 (0.18)	0.25 (0.10)
-Speech errors per clause	0.26 (0.15)	0.20 (0.10)	0.21 (0.12)	0.38 (0.19)	0.28 (0.11)	0.36 (0.20)

Number of children who produced a disguised adaptor

In total, 22 disguised adaptors were observed. The absolute number of disguised adaptors produced by each age group was three times for 9-year-olds, three times for 10-year-olds, twelve times for 11-year-olds, and four times for 12-year-olds. Three 9-year-olds, three 10-year-olds, five 11-year-olds, and two 12-year-olds produced a disguised adaptor at least once during their narrations. There were no 7- or 8-year-olds who produced disguised adaptors. A Fisher's exact test was used to examine the relationship between the age of the group and the number of children who produced disguised adaptors. There was a significant association between them (Fisher's exact test, p = .03). A residual analysis was conducted to find out where the significant differences among age groups were. The analysis indicated that 11-year-olds produced disguised adaptors more often than the other age groups did.



ntto: saisho: sono neko ga: tweety o: sagashi ni: it-te: [de **mi-ta** **mi-**ta nekotan] ttsut-te: sorede [(1)% (2)sono] (2.07) (3)sagashi ni it-ta tokoro wa: inu: toka neko wa dame tte iu omise de

'well, at first, the cat went to look for Tweety and (he) said "(he) saw, saw the pussy cat," then, well the place where (he) went to look for the cat was a shop where dogs and cats were not allowed'

Figure 2: A 11-year-old girl telling a story.

Figure 2 shows a typical case of a disguised adaptor in which an 11-year-old girl described the first of eight scenes, in which *Sylvester the Cat* goes into a building to catch *Tweety Bird* who lives there. Regarding the scene where Tweety said that he saw the cat, "mita mita neko tan ttutte sorede", the girl raises her right arm and puts her fingers into an O-shape ((1) in Figure 2). Given that she uses the right side of the space and a hand shape as if it is holding an object while talking about where Sylvester went to find Tweety, the right hand movement seems to be part of the preparation phase to depict the building that Tweety is in.

However, while saying 'sono', which functions as a filled pause and also means an article "the" in English, she stopped production of the gesture's stroke and modified it into a disguised adaptor of rubbing her right eye with her right hand ((2) in Figure 2). After that, an unfilled pause, which lasted about 2 seconds, occurred until the next word started. During this pause, she retracted the disguised adaptor and described in speech the place where the cat went to find the bird. This case can be interpreted as being that she first tried to depict the building that the cat goes into by using both speech and gestures. However, because she could not remember the proper name for it, she abandoned the gesture in the middle of the production and modified it into a disguised adaptor.

Temporal relationship between speech and disguised adaptor

Each disguised adaptor could be categorized into one of six types (Table 2). The most frequent type was the co-occurrence of a disguised adaptor and intact speech (Type 1), followed by a combination of a disguised adaptor and intact speech after a false start(s) (Type 5) and synchronization of a disguised adaptor and a filled pause after a speech error and/or filled pause (Type 6).

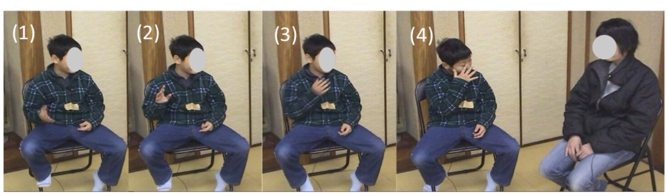
Table 2: Number of cases categorised into temporal relationships

Types [1)]	Absolute number (%)
1. PREP → Intact speech + DA	7 (32)
2. PREP → Filled pause + DA	3 (14)
3. PREP → Unfilled pause + DA	2 (9)
4. PREP → Speech error + DA	1 (5)
5. PREP → Speech error (→Speech error) → Intact speech + DA	5 (23)
6. PREP → Filled pause and/or Speech error → Filled pause + DA	4 (18)
Total	22 (100)

1) Abbreviations and notations: DA (disguised adaptor), PREP (preparation phase), → (order of temporal direction), + (synchronization of elements)

To see whether disguised adaptors were related to word searches, the six types were further classified into two groups in terms of whether the disguised adaptor occurred with a speech error, (un)filled pause, or intact speech element. The former group is termed the *dysfluent speech combination* and includes Types 2, 3, 4, and 6, whereas the latter group is named the *fluent speech combination* and includes Types 1 and 5. The number of children who produced each combination was counted for each age group. Three 9-year-olds, three 10-year-olds, three 11-years-olds, and one 12-year-old produced fluent speech combinations. Five 11-year-olds and one 12-year-old produced dysfluent speech combinations. No children aged 9 or 10 produced dysfluent speech combinations. A Fisher's exact test was used to examine the relationship between age group and each combination. A significant association was found only in the dysfluent speech combination (Fisher's exact test, p = .002). In contrast to the fluent speech combinations produced by 9- to 12-years-olds, dysfluent speech combinations were not produced by 9- and 10-year-olds, and all of 11-year-olds produced dysfluent speech combinations.

Figure 3 show a 9-year-old boy who produced a fluent speech combination. The figure shows him explaining a scene in which Sylvester got inside Tweety's apartment, but was struck by the bird's owner with an umbrella. While he was talking about the umbrella, he raised his right hand and shaped his hand as if it were holding an umbrella ((1) and (2) in Figure 3). However, without finishing the preparation as a gesture, he moved his hand to his eye to scratch it as a disguised adaptor. Unlike the girl in Figure 1, who produced a dysfluent speech combination, his speech did not contain obvious speech errors or pauses. Thus, his disguised adaptor may have been caused by other factors besides a word search (this point is taken up in the discussion section).



sorede ouchi no naka ni hait-ta n dakedo sono kainu[shi no **obaa**chan ga] dete[(1)ki-te (2)kasa de: (3)# (4)nagurare-te: tsugi wa are]

'and (he) got inside the house, the grandma, the owner comes out and (he) is struck with an umbrella, and the next is, umm'

Figure 3: A 9-year-old boy telling a story.

Discussion

The present study investigated disguised adaptors as an index of a child's ability to monitor his or her own gestures by focusing on the relationship between disguised adaptors and speech flow. The results showed that the gesture frequency and the proportion of speech errors increase with age, especially in the late elementary school period, whereas the proportion of unfilled pauses decreases with age. The increase in speech errors suggests that children tend to dedicate much effort to planning coherent narratives especially from the age of 10. Considering pauses may reflect cognitive processes underlying speech planning including word search, syntax, conceptual and articulation planning (Schönpflug, 2008), it is considered that children gradually acquire the ability to plan speech quickly during their elementary school years.

In this study, disguised adaptors were produced by children who were more than 9-years-old. None of the 7- and 8-year-olds used disguised adaptors at all. This result indicates that children become aware of their gestures as an informational resource for listeners from the age of 9 onwards. In other words, they acquire the ability to monitor their gestures from the age of 9.

The analysis of the temporal relationship between a disguised adaptor and speech fluency suggests that disguised adaptors are caused by not only speech dysfluency but also other factors. Children from 9- to 12-years-old produce fluent speech combinations, but only the higher graders produced dysfluent speech combinations. This implies that disguised adaptors of 11- and 12-year-olds are partly caused by the act of searching for an adequate word or planning a sentence. In these cases, children may notice that they have to stop speaking to retrieve a word or re-plan a sentence, and accordingly they modify their gestures to a disguised adaptor in the middle of gesture production. Based on these results, I will discuss why disguised adaptors appear around the age of 9 years in terms of cultural and cognitive factors

Cultural influence

Previous studies on the gestures of elementary-school-age children suggest a cultural influence. Some studies have reported that the frequency of gestures consistently increases during the elementary school years. These trends appear across cultures, although most of the studies were on children in Indo-European language cultures (e.g., Colletta et al., 2014). However, this study on Japanese children showed that the frequency of gestures decreases temporarily in the middle grades compared with in the lower grades or higher grades. This difference may come from their educational environment. In Japan, sometimes pupils are implicitly and explicitly warned by their teacher to avoid fidgeting or moving their hands when the teacher or another child is speaking or sometimes even when they themselves speak. In fact, speakers in Asia sometimes learn not to gesticulate (Neu, 1990). These findings suggest that Japanese children as young as 9-years-old attempt to embody the rule about hand movement during communication. This may be related to why disguised adaptors produced by 9- and 10-year-olds do not synchronize dysfluences. Because they seem to start noting that their hand movements can be read by someone as symbols, even when they do not have a problem with word search, they may try to suppress their hand movements.

Cognitive change

As a factor influencing the emergence of disguised adaptors from the age of 9 or 10, one may consider cognitive changes occurring during this period. Piaget & Inhelder (1969) suggested that the period is considered to be the concrete operational stage at which abstract logical thought is first applied to the physical world. Karmiloff-Smith (1986), who investigated the development of metalinguisitc awareness in 4- and 12-year-olds, found that many children from 9 years onwards explicitly have metalinguistic awareness. Thus, children in the middle grades of elementary school seem to develop metacognitive knowledge to notice that there are underlying rules or mechanisms in the physical world and human communication. At same time, they may also become aware that gestures are informational resources for the listener. Because children in this period are sensitive to such rules, they start monitoring their expressions to check whether the message in the expression is adequate given the communicative context. This awareness seems to result in the emergence of disguised adaptors and an increase of speech errors in children in the later years of elementary school. Ito and Tahara (1985) found that 10-year-olds had poorer usage of the postpositional particle *WA* in comparison with other age groups. The speculated that because children in this period are just beginning to notice and attempt to grasp the multifunctional nature of language devices, their performance seems to decline temporarily. This suggests that the ages of 9 and 10 can be seen as the transitional period in which Japanese children begin noticing the communicative function of gesture and linguistic system, and monitoring them.

Examining when children suppress gestures contributes to an understanding of children's gestural development. Just as certain self repairs in speech that are spontaneously made by children during narratives reflects metalinguisitc awareness that they have acquired (Karmiloff-Smith, 1986), the emergence of disguised adaptors implies that children have an awareness of gestures. This study showed that although the production of gestures may be mostly an unconscious process (Goldin-Meadow & Butcher, 2003), the speaker can notice that she is producing a gesture after she raises her hand for a gesture, and that this awareness begins at about 9 years of age. Future task is exmaning whther disguised adaptors are rubost phenomonen by collectiong more data from other age groups and from other cultuore groups.

References

Beattie, G. (1983). *Talk: An Analysis of Speech and Non-verbal Behaviour in Conversation*. Open University Press: Milton Keynes.

Colletta, J. M., Guidetti, M., Caprci, O., Cristilli, C., Demir, O. E., Kunene-Nicolas, R. N., & Levine, S. (2014). Effects of age and language on co-speech gesture production: an investigation of French, American, and Italian children's narratives. *Journal of Child Language*, 42, 122-145.

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding, *Semiotica,* 1, 49- 98.

Fujii, M. (1997). Effects of self-touching behavior on the performance of lexical Retrieval. *Japanese Journal of Psychology*, 68, 187-196.

Fujii, M. (1999). Effects of gestures in communication: Developmental investigation of speech and gesture. *Japanese Journal of Educational Psychology*, 47, 87-96.

Goldin-Meadow, S., & Butcher, C. (2003). Pointing toward two-word speech in young children, in: S. Kita (Ed.), Pointing: Where Language, Culture, and Cognition Meet (pp.84-107).Mahwah, N.J., Lawrence Erlbaum Associates.

Ito, T., & Tahara, S. (1985). A psycholinguistic approach to the acquisition of multifunctionality in Japanese particles wa and ga. *Descriptive and Applied Linguistics*, 18, 121-131.

Karmiloff-Smith, A. (1986). From meta-processes to conscious access: Evidence from children's metalinguistic and repair data. *Cognition*, 23, 95–147.

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. Key (Ed.), The relationship of verbal and nonverbal communication (pp. 207-227). The Hague, the Netherlands: Mouton.

Mayberry, R. I., & Nicoladis, E. (2000). Gesture reflects language development: Evidence from bilingual children. *Current Directions in Psychological Science*, 9, 192-196.

McNeill, D. (1992). *Hand and Mind*, Chicago: University of Chicago Press.

Neu, J. (1990). Assessing the role of nonverbal communication in the acquisition of communicative competence, in R. C. Scarcella, E. S., Andersen & S. D. Krashen (Eds), Developing Communicative Competence in a Second Language (pp.121-138). New York: Newbury House.

Piaget, J. & Inhelder, B. (1969). *The Psychology of the Child*. New York: Basic Books.

Pine, K. J., Bird, H., & Kirk, E. (2007). The effects of prohibiting gestures on children's lexical retrieval ability. *Developmental Science*, 10, 747–754.

Schönpflug, U. (2008). Pauses in elementary school children's verbatim and gist free recall of a story. *Cognitive Development*, 23, 385-394.

Sekine, K. & Furuyama, N. (2010). Developmental change of discourse cohesion in speech and gestures among Japanese elementary school children. *Rivista di psicolinguistica applicata*, 10, 97-116.

Sottofattori, S. M., & Nicola, F. (2007). Self-Manipulation and the Tourette's syndrome, in Proceedings of the 3rd Conference of the International Society for Gesture Studies, pp.46-47. Chicago, America, June.

Volterra, V., & Iverson, J.M. (1995). When do modality factors affect the course of language acquisition? in: E, Karen, J. S. Reilly (Eds.), Language, Gesture, and Space (pp. 371-391). Hillsdale, N.J., Lawrence Erlbaum Associates.

Wigglesworth, G. (1990). Children's narrative acquisition: A study of some aspects of reference and anaphora. *First Language*, 10, 105-125.