



Bimodal compositions of gesture and vocalization at the beginning of verbal communication / *Composiciones bimodales de gesto y vocalización en el inicio de la comunicación verbal*

Fernando-Gabriel Rodríguez & Silvia-Ana Español

To cite this article: Fernando-Gabriel Rodríguez & Silvia-Ana Español (2016) Bimodal compositions of gesture and vocalization at the beginning of verbal communication / *Composiciones bimodales de gesto y vocalización en el inicio de la comunicación verbal*, *Infancia y Aprendizaje*, 39:4, 661-693, DOI: [10.1080/02103702.2016.1215086](https://doi.org/10.1080/02103702.2016.1215086)

To link to this article: <https://doi.org/10.1080/02103702.2016.1215086>



Published online: 31 Aug 2016.



Submit your article to this journal [↗](#)



Article views: 301



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 2 View citing articles [↗](#)

Bimodal compositions of gesture and vocalization at the beginning of verbal communication / *Composiciones bimodales de gesto y vocalización en el inicio de la comunicación verbal*

Fernando-Gabriel Rodríguez^a and Silvia-Ana Español^b

^aUniversidad Argentina de la Empresa; ^bFlacso – Conicet

(Received 9 August 2015; accepted 11 May 2016)

Abstract: This study contributes empirical evidence regarding the central role that gesture plays in the general process of language acquisition. Previous studies have highlighted this factor during the holophrastic stage (12–18 months), however a void exists in relation to the stage immediately following and immediately preceding the appearance of grammatical structures, the two-word stage. A single case longitudinal study was performed using the constant comparative method. Frequency data were obtained, indicating a permanent choice of the oral communication modality over the gestural modality, while at the same time a prevalence of bimodal compositions (gesture-vocalization) over oral unimodal compositions was seen. This suggests that bimodal compositions function as a cognitive-communicative procedural antecedent of the exclusively oral compositions that will prevail with the arrival of two words before the handling of grammatical structures, in other words, before the (relative) *unimodalization* of verbal expression. These results subscribe to the non-modularistic theory of language from the role of gesture in the language acquisition process.

Keywords: early childhood; gesture; vocalizations; compositions; bimodality

Resumen: Este trabajo aporta evidencia empírica del rol central que cumple el gesto en el proceso general de adquisición del lenguaje. Estudios previos han puesto de relieve este factor durante el período llamado holofrástico (12–18 meses), pero existe un vacío en relación con el período inmediatamente siguiente e inmediatamente previo a la aparición de estructuras gramaticales, el período de dos palabras. Se realizó un estudio longitudinal de caso único empleando el método comparativo constante. Se obtuvieron datos de frecuencia que indican una elección permanente de la modalidad de comunicación oral por sobre la gestual, pero al mismo tiempo una prevalencia de las composiciones bimodales (gesto-vocalización) por sobre las unimodales orales. Ello sugiere que las composiciones bimodales funcionan como antecedente procedimental cognitivo-communicativo de las composiciones exclusivamente orales que prevalecerán con el arribo de los acoplamientos de dos palabras antes del manejo de estructuras gramaticales, esto

English version: pp. 661–676 / *Versión en español*: pp. 677–692

References / *Referencias*: pp. 692–693

Translated from Spanish / *Traducción del español*: Jennifer Martin

Authors' Address / *Correspondencia con los autores*: Fernando Gabriel Rodríguez, Universidad Argentina de la Empresa, Lago Lácar 505, Bella Vista, CP. 1661 Provincia de Buenos Aires, Argentina. E-mail: fgrxyz@gmail.com

es, antes de la *unimodalización* (relativa) de la expresión verbal. Estos resultados abonan la tesis no-modularista del lenguaje a partir del rol del gesto en el proceso de adquisición lingüística.

Palabras clave: infancia temprana; gesto; vocalizaciones; composiciones; bimodalidad

Verbal communication emerges by the age of 12 months in the form of single words. This holophrastic stage then extends up until around 18 months of age. It is thus named because these words are susceptible to acquiring quasi-propositional meanings (the child says ‘water’ with volitional meaning — ‘I want water’ — or informative meaning — ‘there is water there’ — and the pragmatic-contextual coordinates allow interlocutors to understand the meaning). The first compositions of word pairs then appear, mostly made up of descriptive or referential terms without the presence of logical connective elements. This stage, called the *two-word* stage or *first compositions* stage, extends from approximately 19–27 months of age, after which verbalizations acquire a larger extension and some morphosyntactic structures appear (Serra, Serrat, Solé, Bel, & Aparici, 2000).

During the holophrastic production stage, gesture and word work together to form a message in which the contents of the two communicative channels mutually support each other (Butcher & Goldin-Meadow, 2000; Capirci & Volterra, 2008). This reveals that a child can adjoin two diverse communicative units before he is able to achieve this with two verbal units in the two-word stage. Three diverse forms of connecting gesture and word have been distinguished: equivalence, complementarity and supplementarity (Volterra, Caselli, Capirci, & Pizzuto, 2005). In the first, the two units carry the same meaning towards a single referent (a pointing gesture coinciding with a verbal deictic: ‘this’). In the second, gesture and word continue to be coreferential but they differ in the amount and quality of information that they provide (pointing out a small car and saying ‘car’, where the second term categorizes the object indicated, with a conceptual term mediating between the object and the gestural movement). Finally, in the third, the coupling of gesture and word is integrated by two units with distinct referents: the utterance ‘mama’ coupled with pointing to a shoe would indicate some connection between the two ideas *mama* and *shoe* (= *mama’s shoe*).

In the holophrastic stage the rate of messages with gesture-word supplementarity increases as equivalence and complementarity rates decrease (Butcher & Goldin-Meadow, 2000; Iverson, Capirci, & Caselli, 1994). The co-occurrence of gesture and word provides, in cases of supplementarity, a guarantee that a step prior to the possibilities of expressing two different contents only in the verbal modality exists in the development of a child’s communicative abilities. It is feasible to believe that, in supplementary compositions, gesture operates as a precedent for the subsequent juxtaposition of words in the two-word stage, and therefore as the *primary* form under which the human subject communicates two different representations/concepts in a connected way. In turn, the equivalent-complementary-supplementary progression supports the gradualist hypothesis of

language acquisition, which suggests that there is continuity between preverbal communications and the syntactically coordinated coupling of words.

Ontogenetic studies on relationships between gestures and words tend to only go up to the age of 18/20 months, which is the end of the holophrastic stage. Even though the study of adult communication has shown the effectiveness of a simultaneous approach to both modalities (McNeill, 1992), the joint study of gesticulation and pregrammatical verbalizations during the two-word stage has still not been suitably explored. This stage is especially attractive because a crucial transformation within communicative development is produced: gestural and oral modalities are no longer composed *on an equal footing* and gesture begins to accompany word in a different way. Thus, words enter into the terrain of grammaticalization and gestures (made up of non-articular characteristics) become marginalized. Gesture is a *Gestalt* inasmuch as it cannot be split up into lesser significant elements (corresponding to the morphemes of a word). Furthermore, it is non-articular because even though it is susceptible to linking with other signs, its linking does not respect combination rules (which characterizes the grammatical articulation process). The course, mode and temporality of this fickleness in communication still await a detailed observation. As two well-known authors in this field have written on this matter, ‘the goal is now to explore and better describe the transition through which children’s gestures become organized into the adult system of speech and gesture’ (Capirci & Volterra, 2008, p. 37).

This approach stands before modularity positions (Chomsky, 1984; Fodor, 1983) in which the faculty of language constitutes an impenetrable mental processing device, only operational for a certain type of representation. Studies focused on the progress of expressive skills in their entire spectrum, unlike those studies restricted to pure linguistic acquisition, have mostly been against an innate modularity, occasionally permeable to the possibility of a gradual modularization (Karmiloff-Smith, 1995). In this way language emerges from the cooperative use of the child’s previous resources: joint attention, goal intentionality, mindreading, intentional communication of a single sign, gesture-word bimodal coupling and word-word unimodal coupling. These pregrammatical skills finally arrive at the grammatical articulation of verbal words.

Given the scarce information on bimodal communicative skills from the two-word stage, this study sought to contribute evidence over the isolated and coordinated use of gestures and vocalizations during the two-word stage from a developmental single case, longitudinal study. The joint approach of more than one modality reclaims Piaget’s idea of a general semiotic function, opposed to modularist approaches in which language and gesticulation would be unconnected processing systems. Understanding development as a *dynamic system* implies conceiving it as subject to constant movement and a progression that also includes impasses and setbacks (non-linear progression) that are multicausal and not reducible to a causal relationship of the ‘A then B’ type. The single case design enables a detailed monitoring of a subject and with this, the possibility of looking at intra-subject variations, which is very valuable to a dynamic conception for allowing change appreciation in a dimension unseen in cross-sectional studies.

Method

Participants

An adult-child dyad was observed. The single subject was an Argentine boy from the city of Buenos Aires, who was born full-term and who had no problems or illnesses during development. He was from a middle-class home, with a brother three years older and professional parents, and a member of a Spanish-speaking monolingual family. In fact, the subject was the first author's youngest son. The child's age in the first session, according to Piaget convention [year; month (day)], was: 1; 7 (13); in the last session: 2; 5 (1). The adult's role in the dyad was alternately carried out by the mother, the father-researcher and a nanny. This way the reality of the child's communicative skills was reflected free of biases that could result from a particular interlocutor.

Procedure

Fifteen sessions of adult-child interactions that occurred in the home and within a natural context were recorded. Care was taken during the filming so that the bodies of both the child and adult were always in the shot, especially their hands, face and torso, so as to not miss any movements or expressions from either of them. A total of seven hours and 30 minutes were filmed from 15 sessions of 30 minutes each occurring at 20-day intervals. The recorded situations occurred freely and developed within a playful atmosphere consistent with age (give and take games, reading illustrated stories, assembling and insertion games, etc.). In all cases, the adults initiated the exchanges by proposing activities that were then allowed to progress according to the child's spontaneity. The filming was alternately performed by the father-researcher or by the mother.

In each videotaped session the child's communicative interventions were identified and counted. Communication was considered to be every intentional and signifying interaction (Rivière, 1998), that is: all actions meant to contact the interlocutor through the use of signs. Cases in which the child was self-interacting (self-gesture, an acknowledging verbalization) were excluded because, although they are relevant for explaining cognitive development, they are outside of our definition of communication. An intervention was considered to be any communicative participation from the child in the interactions with the adult partner (gestures, vocalizations and compositions).

An observation code was created according to the constant comparative method (Strauss & Corbin, 2002). Standard categories were first adopted for the study of gestures and verbalizations during infancy and early childhood. Gestures were organized according to the classic division between deictic and representational gestures, defined in the following section; oral expression contemplated both lexical and non-lexical vocalizations; and for compositions, the aforementioned categories of equivalent, complementary and supplementary were adopted (Volterra et al., 2005), considering unimodal (gesture-gesture, vocalization-vocalization) or bimodal (gesture-vocalization) compositions. In a second phase, with the aim of exploring the

entire variety of the child's communicative resources, new categories were devised for interventions with classifications that were impossible in the standard categories of the first code. Gestures and vocalizations were classified as *pure gestures* or *pure vocalizations* when they were performed or uttered, respectively, in isolation (without the accompaniment of another gesture or vocalization), or as *compositions* when they were joined with other units (gesture or vocalization, word or non-word, in varying number). All expressions of *simultaneous* signs (gesture-vocalization, gesture-gesture [since hands and face can indicate at the same time]) or *successive* signs (vocalization-vocalization or gesture-vocalization) that, without a pause of more than a second, enabled understanding of its semantic integrity by context (the mutual belonging of signs as parts of a message) were considered to be compositions. Compositions were unimodal when their composing units only corresponded to one modality (gestural/oral) and bimodal when they placed both modalities into relation for a single emission.

The observation code was applied to each one of the sessions. Each session's content was independently classified by two researchers (the authors of this article). The differently classified cases were discussed and contended. For the few instances in which an agreement was not reached, the intervention in question was discarded. This procedure was consistent with that utilized in social interaction studies (Forster & Iacono, 2014; Jonsson et al., 2001).

The code was formed by using the following categories and subcategories.

Gestures

Gestures encompassed the following types: (1) *deictic* (*pointing, showing, reaching* and *request; giving*, in the limit with action, was not counted since *doing* and proposing at the same time is not purely and strictly a sign); (2) *representational* (total/partial mimicry of an object or action: imitation of an airplane by stretching arms out to sides); (3) *connective* (a connection for other signs: marking of both times for any before/after episode by moving the hand from one side to the other); (4) *negation/affirmation* (shaking the head to indicate yes or no); (5) *expressive* (gives the emotional tone in order to understand another communicative element: a grimacing face, a smack on some surface to highlight a verbal expression); and (6) *arbitrary* (hand waving to indicate 'goodbye'). The first two types, along with the last one, corresponded to a traditional categorization (for example, Iverson & Thal, 1998); the third to McNeill (1992); and the fourth and fifth were classificatory contributions from this study.

Vocalizations

Vocalizations included words as well as oral expressions outside of lexicon (non-words) in which there was recognition of communicative intentions and aptitude for integrating them into compound meanings. Non-words — expressive interjections ('eh!', 'oh!', 'ay!') and non-expressive modulations (e.g., 'ta', or else 'ta-ta-ta') effectively emitted for interacting — were reevaluated in studies on language acquisition (Karousou &

López Ornat, 2013), recognizing in them a predictive value even greater than that of gestures in future linguistic skills. With non-expressive modulations, which could be monosyllabic or multisyllabic, the first were counted as nothing more than units, whereas the multisyllabic modulations were divided into simple syllabic units, except when they presented a constant form (trisyllabic unit *ta-ta-ta* temporally separated — for more than a second — from other utterances) and were used in at least two sessions within semantically defined contexts (thus meeting the standardized definition of protowords).

Compositions

Compositions were classified according to (a) type and amount of components and (b) the semantic collaboration between the components. Classification (a) distinguished the following compositions: gesture-gesture (G-G), gesture-vocalization (G-V), vocalization-vocalization (2V), gesture-gesture-vocalization (2G-V), gesture-vocalization-vocalization (G-2V), gesture-gesture-vocalization-vocalization (2G-2V) and unimodal compositions of three, four or more vocalizations. Classification (b) differentiated the compositions into *equivalent* (semantically identical units), *complementary* (semantically convergent), *supplementary* (semantically different), *emotive* (where the gesticulation or the vocalization simply contributed an emphasis on the meaning of another constituent and not, as in the previous categories, informative content — for example, the name of a toy accompanied by the expressive gesture of surprise upon its discovery) — and *doubtful* (when a decision could not be made as to which of the other four categories a particular composition should be assigned to). Emotive compositions are normally not covered in studies of this type; however, having recognized their importance, they were incorporated into the second aforementioned phase of the coding process.

The reduplicative non-word vocalizations of the ‘ta-ta-ta’ type were not counted as part of the compositions since they were considered to be simply a vocalic exercise (sometimes quite long), but they were counted among the oral emissions, as mentioned above. Only the interjections and protowords were considered to be vocalizations constituent of compositions.

The compositions were limited by their semantic unit and by their relative temporal coincidence, as already explained. With regard to the first, the intervening signs in the composition had to have the same reference (pointing + verbalization ‘here’ = G-V) or link two referents (‘me’ + ‘ball’ = 2V) or proclaim a referent’s attribute (pointing + ‘this’ + ‘large’ = G-2V) or express an emotion with regard to a referent (‘not here’ + gesture of surprise = G-2V). The temporal pattern involved coincidence but not perfect synchrony (an impossible criterion if wanting to take compositions with two or more vocalizations into account, which necessarily involves a successive arrangement of signifying elements, since they belong to the same modality). Patterns considered to be the end of a gestural composition were the absolute change in semantic referentiality and a temporal gap longer than a second. The order of the composition’s elements was not

considered relevant given that in the stage studied they were understood as juxtapositions without clear evidence of grammaticality (López Ornat, 1999). Thus, the compositions were not viewed as ordered pairs (or triples, etc.) but simply as signalling couplings.

Results

Gestures and vocalizations

Table 1 shows the frequency of the gestures and vocalizations in their different types of occurrences (pure and in composition) counted during the observation period. Throughout the period studied a clear predominance of interventions with pure elements (isolated gestures or vocalizations) was recorded. These occurred more than compositions, with a difference of 56.3% in session 1 and of 20.6% in session 15 (on average a difference of 33.8%). The compositions (gesture-gesture, gesture-vocalization, vocalization-vocalization) never surpassed single element interventions.

Table 2 shows that the frequency of vocalizations (pure, in composition with gesture or vocalization, and totals) was higher than that of gestures (pure, in composition with vocalizations or with other gestures, and totals) throughout the whole period. This indicated a general preference for the oral modality in the interventions from each one of the sessions. A positive correlation was identified between total gestures and total vocalizations ($r_s = .554$; $p < .05$), which would show that the behaviour of one or another modality, beyond the predominance of oral expression, was maintained in proportion and that the gradual increase of the oral signs did not occur at the expense of gestural signs.

For the gestural modality, the gestures integrated into compositions constituted 85.2% of the total gestures, with a difference in the frequency average throughout the 15 sessions of 62.2 above the frequency of pure gestures average. A slight increase was seen in pure gestures towards the second half of the period studied, which from session 8 onwards did not go below two digits and on three occasions had frequencies higher than 20 (sessions 10, 13 and 14).

For vocalizations, and contrary to what occurred with gestures, a predominance of pure interventions over vocalizations integrated into compositions was observed. This reversed in the last session, where vocalizations in composition, after obtaining frequencies very similar to pure vocalizations during the previous sessions (12, 13, and 14), reached a higher frequency in a ratio of 246–170.

The comparison between pure gestures and pure vocalizations produced an overwhelmingly higher result for the oral modality: during the 15 observation sessions, the average frequency of pure vocalizations was higher than the average frequency of pure gestures by 200.7 (see Table 2). Comparison between the gestures integrated and the vocalizations integrated into compositions also revealed a difference of 95.5 in the average of frequencies, favourable to the latter. Towards the second half of the period the difference in values intensified and the vocalizations-compositions exceeded 100 in most of these sessions (except sessions 10 and 11).

Table 1. Frequency of interventions with pure elements (gestures or vocalizations) and compositions.

Session	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Totals
Total interventions	403	238	238	313	339	245	311	421	358	229	350	465	455	432	310	5,105
Interventions with Pure Elements	315	175	137	214	192	163	226	264	241	166	248	308	298	286	187	3,420
Interventions with Pure Elements %	78.1	73.5	57.6	68.3	56.6	66.5	72.7	62.7	67.3	72.5	70.9	66.2	65.5	66.2	60.3	66.9
Interventions with Compositions	88	63	101	99	147	82	85	157	117	63	102	157	157	146	123	1,687
Interventions with Compositions %	21.8	26.5	42.4	31.6	43.4	33.5	27.3	37.3	32.7	27.5	29.1	33.8	34.5	33.8	39.7	33.0

Table 2. Frequency of pure gestures, pure vocalizations and compositions.

Session	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Totals
Total Gestures	68	62	99	98	143	57	59	114	81	73	92	125	111	121	87	1,390
Pure Gestures	7	8	2	13	14	6	0	12	18	27	18	16	26	21	17	205
Composition Gestures (gesture-gesture, gesture-voc.)	61	54	97	85	129	51	59	102	63	46	74	109	85	100	70	1,185
% of Composition Gestures / Total Gestures	89.7	87.1	98	86.8	92.3	89.5	100	89.5	77.8	63.1	80.5	87.2	76.6	82.7	90	85.2
Vocaliz. Totals	455	245	247	326	367	295	353	491	421	234	380	538	537	527	416	5,832
Pure Vocalizations	308	167	135	201	178	157	226	252	223	139	230	292	272	265	170	3,215
Compos. Vocalizations (gesture-voc., voc.-voc.)	147	78	112	125	189	138	127	239	198	95	150	246	265	262	246	2,617
% Vocaliz. Compositions /Voc. Totals	32.3	31.8	45.3	38.3	51.4	46.8	35.9	48.6	47	40.6	39.5	45.7	49.3	49.7	59.1	44.8

Compositions

Compositions according to the type and quantity of components

Table 3 shows the compositions found according to type (gesture or vocalization) and quantity of components in each one. Throughout the period the predominant composition type was that of gesture-vocalization (G-V = 53.9%). Adding up the frequencies of all the types of bimodal emissions (G-V, 2G-V, 2V-G, 2G-2V and 3V-G), this result was higher than that of the total frequency of unimodal composition cases (2V, 3V, and 4V, or G-G, these having $n = 4$). Bimodal compositions numbered 1,152 (68.3% of total compositions from the subject) versus 531 compositions with only vocalizations (31.5% of total compositions). The remaining 0.23% corresponded to cases of composition without vocalization (G-G).

Spearman's rank correlation coefficient was used to identify the following correlations: (a) unimodal, between 2V and 3V ($r_s = .523$; $p < .05$); and (b) bimodals, between 2V-G and 3V-G ($r_s = .729$; $p < .01$) and between 3V and 3V-G ($r_s = .747$; $p < .01$). Overall, these results would indicate that the growth pattern for the quantity of components did not only occur in the oral modal but that the bimodal compositions, with gestures, far from becoming recessive, accompanied the increase in the child's communications.

According to the Observation Code, the vocalization designation covers both words and non-words. Because words contribute decisively to the grammaticalization process, the word-gesture category was distinguished as a subset within gesture-vocalization compositions and the word-word category was considered to be a subset within vocalization-vocalization compositions. This distinction was important for counting the lexical units integrated into compositions and for closely following these compositions in their progress towards *unimodalization* (verbal) over which grammar is developed. The frequencies of these two variables are detailed in Table 4. According to the expected course of the child's linguistic development, the word-word compositions increased with the passing of time, but not at the expense of the gesture-word compositions.

Table 3. Frequency of compositions according to the type and quantity of components (G/gesture; V/vocalization).

Session	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Totals	%
G-V	44	50	90	77	112	40	47	77	44	35	59	81	69	55	30	910	53.9
G-G	0	0	1	0	0	0	0	1	1	0	0	0	0	0	1	4	0.2
2G-V	1	0	0	1	1	0	0	2	0	1	1	2	0	3	1	13	0.8
2V	16	8	3	11	13	18	23	54	47	15	24	40	58	42	35	407	24.1
2V-G	11	2	5	4	13	10	11	18	15	7	11	18	11	17	26	179	10.6
2G-2V	0	1	0	1	0	0	0	0	0	1	1	2	1	3	2	12	0.7
3V	12	2	2	5	6	13	3	3	8	3	6	11	13	8	17	112	6.6
3V-G	4	0	0	0	2	1	1	1	2	0	0	2	3	16	6	38	2.2
4V	0	0	0	0	0	0	0	1	0	1	0	1	2	2	5	12	0.7

Table 4. Frequency of gesture-word and word-word compositions compared to frequency gesture-vocalization and vocalization-vocalization. (G/gesture; V/vocalization; W/word)¹.

Session	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Totals
G-V	60	53	95	83	128	51	59	98	61	44	72	105	84	94	65	1,152
G-W	51	41	93	69	91	46	50	91	59	34	69	92	63	82	59	990
V-V	28	10	5	16	19	31	26	58	55	19	30	59	73	52	57	529
W-W	0	0	0	5	5	4	10	20	38	7	16	32	47	34	42	260

Note: ¹This table accounts for, unlike Table 3, all G-V and V-V compositions regardless of the quantity of each modality’s components. So here G-V encompasses G-V, 2G-V, 2V-G, 2G-2V and 3G-V from Table 3, while V-V encompasses 2V, 3V and 4V from Table 3.

Compositions according to the type of collaboration between the components

Whether the gesture and vocalization components integrated into equivalent (CEQ), complementary (CC), supplementary (CS), emotive (CE) or doubtful (CD) compositions was assessed. Table 5 shows the frequencies of each type identified.

A higher frequency of equivalent compositions was observed, which were more pronounced in the second half of the sessions. In parallel, although supplementary compositions were found to be present from the beginning, they did not show a progressive increase. One interesting result turned out to be the high frequency relative to emotive compositions (12.5%), which were considered for the first time in this study. Their frequency was greater than that of supplementary compositions (9.5%), which should have predominated according to previous studies, and were slightly less than complementary compositions (14.3%).

Table 5. Frequency of compositions according to the type of collaboration between the components.

Session	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Totals	%
CEQ	48	60	27	61	39	49	49	110	91	49	77	78	74	81	63	956	56.6
CC	8	22	14	41	22	21	15	11	8	6	1	10	30	11	21	241	14.3
CS	9	7	5	20	9	0	6	8	8	0	14	24	22	17	11	160	9.5
CE	20	5	8	9	20	12	6	22	6	1	5	39	17	26	15	211	12.5
CD	3	7	9	16	9	0	9	6	4	7	5	6	14	11	13	119	7.0

Discussion

Gestures and vocalizations

During the entire period studied a greater frequency of pure vocalizations than of gestures and compositions can be observed (see Table 2). This indicates a constant preference for the oral modality, which is in line with the fact that this modality is also the first option with which the child expresses his needs and emotions. Nevertheless,

the correlation between total gestures and total vocalizations indicates that the behaviour of the two communicative modalities is comparable independently of the diverse frequencies. The increase of the child's oral communication in the course of a certain session is accompanied by an increase in gestural communication. It is as if, in his predisposition to interact, the child will use both resources with a strong preference for oral communication, but without neglecting the alternative use of gestures, maintaining a constant proportion between both.

Despite having studied the two-word stage, pure vocalizations are still the dominant expressions. This means that the name 'two-word' only indicates the *occurrence* of verbal element pairs without alluding to its frequency. Nevertheless, in the second half of the period a significant increase of word-word compositions is seen (see Table 4). This reveals that, against a backdrop of invariably dominant vocalizations, compositions of word units, necessary for grammatical verbal communication, increase along with the time variable as a progressive specialization within the overall possibilities of oral expression.

The inverse relationship regarding the behaviour between word-word compositions and pure oral elements suggests the accommodation of different expressive resources into the subject's normal development, whose future holds articulated language, for which the establishment of oral compositions over which grammatical articulation rules can operate is necessary.

Pure gestures, elements that were barely present in the beginning, increased in the final sessions. This may have been due to a general increase in the child's communication rate, according to which, and along with a bimodal explanation, the simultaneous growth of the frequencies for gesture and orality is easy to understand. This interpretation is consistent with the correlation between total gestures and total vocalizations mentioned earlier.

Compositions

A clear predominance of the gesture-vocalization type was attested. This is an interesting result since it highlights that, even in the two-word stage, bimodality dominates the compositions. Thus, the predominance of oral expression elements within the child's total interventions is not accompanied by equal predominance of exclusively oral compositions. This shows that orality is used in most of the interventions either alone or linked to gesture, but not linked (yet) to another oral element.

The growth of the exclusively oral compositions from the middle of the period (see Table 2) is consistent with language progress. If, at the same time, the gesture-vocalization compositions are considered dominant during the entire period, the cognitive transformation aimed at grammatical oral expression could be explained, at this level, as a dynamic process concerning the gradual incorporation of elements from one particular modality into a pre-existing bimodal processing structure. This process is dynamic once the growth of oral unimodal cases does not replace the first composition form made up of two elements, but rather constitutes a nonsubstitutive acquisition. In fact, words will always have support in this bimodal structure, and a permanent communicative aid in gesture

(McNeill, 1992). The connection and mutual assistance between gestural and oral expression points in favour of a processing device that is open to both types of communicative units and is non-encapsulated, as the modularist version of language is intended.

The growth of word-word compositions, sustained but irregular, does not exceed the gesture-word compositions (see Table 4) and is not affected by its relatively stable frequency. This shows that the increase of unimodal compositions of the oral type simply occurs because the oral modality is a favoured option in the communicative interactions of humans.

Frequency of compositions according to the type and quantity of components

Gesture-word compositions are observed from the beginning of the period, but there is no presence of word-word compositions. The first example of this last type is recorded in session 4 ($n = 5$), and they grow in an irregular but constant manner from there. The gesture-gesture compositions barely amount to four throughout the entire period, which leads to the belief that gestural expression is easily produced in pure format or in composition with vocalizations (a matter that displays the point at which oral modality is a pivot for complex expressions).

From the beginning there are also oral modality interventions of more than two vocalizations (3V and 3V-G, considering vocalizations of all type, word and non-word). These data, contrasted with the absence of word-word compositions, show that in principle, orality acts as the expression of non-lexical elements with lexical elements, and that verbal communication *stricto sensu* is only channelled later.

Compositions formed with two unimodal and bimodal elements amount to 78.3% of the total compositions (see Table 3), which clearly reveals that the communicative skills from the period are comfortable with the intermodal pattern of two filler slots or spaces. The ability to connect two elements is consolidated during this period and is projected in parallel as a platform for other compositions of more unimodal as well as bimodal elements.

The consideration of the correlations between different compositions with a number equal or close to the amount of component elements (2V/3V, 2V-G/3V-G, 3V/3V-G), along with the incipient growth of 2V-2G, 3 V-G and 4V compositions produced from session 12 on, suggests that the increase in the amount of components (four) occurs independently from modality. All this credits the idea that the child's communication is interchangeably supported in all the resources (channels) at his disposal and that the relative empowerment of the word within the grammatical system of a language is the product of a previous intermodal collaboration. Even though this point should be placed in parentheses due to the low frequency of some of the compositions (only 12 composed of four elements during the entire period studied), the child's interventions actually show a similar growth, in general, in the number of its elements regardless of their modality.

If we take into account that the set of word compositions shows an upward trend, the general progress of the child's observed communicative skills seems to

be consistent with the general expectation of language development, which means a progressive expansion of vocabulary and compositions of its terms.

Frequencies of compositions according to the type of collaboration between components

Regarding the type of semantic collaboration between the components of a communication, the compositions with the greatest frequency occur in the equivalent category, where both elements (here it was inevitably gesture and word) refer to the same thing and carry the same meaning (gestural and pronominal deixis: pointing + 'this'). Its predominance extends throughout the entire period for a total of 56.6% of the interventions in this category. This must be emphasized because previous studies (Volterra et al., 2005) reported that at the end of the holophrastic stage the supplementary compositions had exceeded the complementary and equivalent compositions. If these results described a reasonable progression from lesser to greater difficulty in the processing of information, passing first from equivalent compositions to a second variety of complementary compositions (a referent diversely alluded to through two units with different meanings and through different modalities, gesturally and orally) and then to a supplementary composition integrating semantically autonomous content (two referents that fulfilled *avant la lettre* the functions of subject and predicate), then the results of our study are removed from such a 'logical' sequence.

It is also the case that the complementary compositions decrease without ever having exceeded the frequency of the supplementary compositions (to which the predominant spot should later be ceded, Volterra et al., 2005). The supplementary compositions, for their part, are present from the very beginning, but there is no sustained upwards trend and at the end of the sessions their frequency is still lower than that of equivalent compositions. These concrete data once again challenge the transition between equivalent — complementary — supplementary compositions. The results obtained in a case study such as ours could suggest that the prevalence of the equivalent compositions is only a matter of communicative style; even so, that sequence deserves a more in-depth consideration as well as new studies.

The lack of a predominant role for supplementary compositions is also unexpected if it is considered, in a novel way, that we have opted not only to include the gesture-vocalization couplings, but also those that are exclusively oral, which have not been studied in previous research. This variant invited us to believe that the predominance of supplementary compositions would be even greater, given that word-word compositions would be added on to the higher frequency of bimodal gesture-word supplementary compositions. However, this more comprehensive definition of supplementary compositions did not result in a greater frequency.

The frequency of emotive compositions, considered for the first time in this study, in relation to the rest (12.5%) suggests that it would be advisable, hereinafter, for studies on gestures and first compositions of gesture-word to include

them. It is possible that this oversight in previous studies had to do with the studies being greatly shaped by interests in linguistic acquisition, with the exclusionary collateral bias of non-referential emotive manifestations.

The persistence, throughout the entire period, of emotive compositions that invariably include some emotional expression (interjection, emotive gesture), along with equivalent compositions, confronted by the growth of supplementary compositions that involve quasi-propositional couplings, suggests that the more simple and original uses of orality (expressed by emotive compositions, which manifest moods) and the uses of more elemental compositions live side by side with the growing number of compositions requiring the handling of operations of greater cognitive complexity (such as the subject-predicate utterance). Thus, in the process of developing communicative tools, the most primitive live side by side with the new strategies (as can be seen in adults, for whom lexical dominance does not cancel out interjections). This sustainability of the first communicative uses of orality together with ulterior ones allow us to frame the development of semiotic skills as a dynamic process of progress and conservation, where what is new is always in constant connection with what is previous and both cognitive strategies interpenetrate (the lexicon expresses emotions and the interjections come under formally correct linguistic structures).

General considerations

This study has found a preference for the oral modality, in pure (one single element) interventions, and a preference for bimodal compositions over unimodal compositions. If bimodality is the norm during the holophrastic stage, the novelty value of our study is that this seems to continue into the two-word stage. The pivot that the gesture-word construction offers to the child's intention to interact does not recede with the increase of compositions that are exclusively formed by words, which would develop in parallel as a second route. Bimodality present in communication from this period, far from being incidental, would credit the cognitive-communicative structure proposed by McNeill (1992).

Regarding our study's limitations, it is important to highlight that, as a single case study, it certainly did not allow generalizations or background explanations, and that the evidence from the longitudinal study is not free from the risk of reflecting the bias of the subjects' personal style, individual maturational features and the non-controlled effect of external variables. Nevertheless, the results of a study of this type are, in the matter that concerns us, 'indispensable' (López Ornat, 1994, p. 8) because they reflect the development process of the phenomenon studied in a detailed way, looking at the progress just as much as at the stopping points and even the setbacks. Future studies must determine if the disorganization of the sequence of equivalent-complementary-supplementary compositions was an idiosyncratic occurrence from the case study or perhaps answers to some regularity. Another of the study's limitations comes from bypassing a qualitative description of the processes referred to. Another study that supplies this type of information is currently in preparation.

Finally, the present study, concentrated in compositions and their connection with what is called pure occurrences, does not provide information about the frequencies of the types of gestures or words that were presented in the period. This information is material for another article related to the present one.

The multimodal approach of early childhood, and of the ontogenetic development of the communicative skills of children in the process of language acquisition in particular, reveal that the intervening factors are of a broader spectrum than that manifested in the limited approaches to verbal progress. Remodelling this process in a multimodal manner involves not only appreciating the spontaneous uses that children perform of their natural expressive resources from a different viewpoint, but also the possibility of providing polyfaceted tools to professionals in the field of education, whose work strategies will become enriched from the contribution of a new perspective. Early clinical and psychopedagogical intervention programmes must re-examine their guidelines and reassign a broader space to gesture and gesture-word compositions, considering them as a central support in language acquisition, and in turn as a process that is not linear nor modular, but on the contrary, as a development subject to accelerations and detentions and open to distinct qualitatively diverse symbolic representations (due to its different expressive vehicle or modal variety, gesture or vocalization). We believe, ultimately, that the present study — as an alternative to thinking of language as a stand-alone device impermeable to other cognitive influences — allows the development of the child's semiotic skills to be conceived as a full-fledged dynamic system.

Composiciones bimodales de gesto y vocalización en el inicio de la comunicación verbal

La comunicación verbal emerge hacia los 12 meses de edad bajo la forma de palabras sueltas. Hasta alrededor de los 18 meses se extiende el período holofrástico, así denominado porque estas palabras son susceptibles de adquirir significados cuasi-proposicionales (el niño dice ‘agua’ con sentido volitivo — ‘quiero agua’ — o informativo — ‘ahí hay agua’ — y las coordenadas pragmático-contextuales permiten a los interlocutores comprender el sentido). Luego surgen las primeras composiciones de pares de palabras conformados mayormente por términos descriptivos o referenciales sin presencia de elementos conectivos o lógicos. Este período, llamado de *dos palabras* o de *primeras composiciones*, se extiende aproximadamente entre los 19 y los 27 meses, cuando las verbalizaciones adquieren otra extensión y aparecen algunas estructuras de morfosintaxis (Serra, Serrat, Solé, Bel, & Aparici, 2000).

Durante la etapa de producción holofrástica, gesto y palabra cooperan en la constitución de un mensaje donde los contenidos de los dos canales comunicativos se apuntalan recíprocamente (Butcher & Goldin-Meadow, 2000; Capirci & Volterra, 2008). Ello revela que el niño puede anexar dos unidades comunicativas de diversa condición antes de que en la etapa de dos palabras lo consiga con dos unidades del código verbal. Se han distinguido tres diversas formas del acoplamiento de gesto y palabra: equivalencia, complementariedad y suplementariedad (Volterra, Caselli, Capirci, & Pizzuto, 2005). En la primera, las dos unidades portan un mismo significado relativo a un mismo referente (gesto de *pointing* coincidiendo con un deíctico verbal: ‘esto’). En la segunda, gesto y palabra siguen siendo correferenciales, pero difieren en la cantidad y calidad de información que brindan (señalamiento de un autito y locución ‘autito’, donde el segundo término categoriza el objeto señalado, llevando a mediar entre el objeto y la ostensión gestual un término de rango conceptual). Finalmente, en la tercera, el acoplamiento de gesto y palabra está integrado por dos unidades con distintos referentes: la emisión ‘mamá’ acoplada al *pointing* de un zapato indicaría alguna vinculación entre las dos ideas *mamá* y *zapato* (= zapato de mamá).

En el período holofrástico la tasa de mensajes con suplementariedad gesto-palabra aumenta conforme disminuye la tasa de equivalencia y de complementariedad (Butcher & Goldin-Meadow, 2000; Iverson, Capirci, & Caselli, 1994). La co-ocurrencia de gesto y palabra proporciona, con los casos de suplementariedad, garantía de que existe en el desarrollo de las habilidades comunicativas del niño un paso previo a las posibilidades de expresar dos contenidos diferentes solamente en la modalidad verbal. Es factible pensar que el gesto opera, en su composición suplementaria, como un precedente de la ulterior yuxtaposición de palabras en el

período de dos palabras, y por ende como la forma *primordial* bajo la que el sujeto humano comunica dos distintas representaciones/conceptos de modo enlazado. A su vez, la progresión equivalente — complementario — suplementario respalda la hipótesis gradualista de adquisición del lenguaje, que plantea una continuidad entre las comunicaciones preverbales y el acoplamiento de palabras sintácticamente coordinadas.

Los estudios ontogenéticos sobre las relaciones entre el gesto y la palabra suelen frenarse hacia los 18/20 meses, al cabo del estadio holofrástico. Aunque el estudio de la comunicación adulta ha mostrado la efectividad de un abordaje simultáneo de las dos modalidades gestual y verbal (McNeill, 1992), el estudio conjunto de la gesticulación y de las verbalizaciones pregramaticales durante la etapa de dos palabras no ha sido todavía explorado convenientemente. Este período es especialmente atractivo porque en él se produce una transformación crucial dentro del desarrollo comunicativo: las modalidades gestual y oral dejan de componerse *en pie de igualdad* y el gesto comienza a acompañar a la palabra de un modo distinto: la palabra ingresa en el terreno de la gramaticalización y el gesto, de características giestálticas no-articulares, resulta marginado. El gesto es una *Gestalt* en la medida en que no puede descomponerse en elementos significativos menores (correspondientes a los morfemas de la palabra). Asimismo, es no-articular porque si bien es susceptible de encadenarse con otros signos, su encadenamiento no respeta reglas de combinación (lo que caracteriza al proceso de articulación gramatical). El curso, modo y temporalidad de esta mudanza en las maneras de comunicar aguarda todavía una observación detallada. Como han escrito dos autoras de reconocida trayectoria en la materia, ‘la meta es ahora explorar y describir mejor la transición a través de la cual los gestos de los niños resultan organizados en el sistema adulto de habla y gesto’ (Capirci & Volterra, 2008, p. 37).

Este planteamiento se enfrenta a las posiciones de corte modularista (Chomsky, 1984; Fodor, 1983) en las que la facultad del lenguaje constituye un dispositivo de procesamiento mental impenetrable, únicamente operativo para representaciones de un determinado tipo. Las investigaciones que enfocan el progreso de las habilidades expresivas en todo su espectro, a diferencia de aquellas restringidas a la pura adquisición lingüística, se han pronunciado mayormente en contra de una modularidad innata, eventualmente permeables a la posibilidad de una modularización paulatina (Karmiloff-Smith, 1995). El lenguaje resulta de esta manera un emergente del uso cooperativo de anteriores recursos del niño: atención conjunta, intencionalidad de meta, lectura de mente, comunicación intencionada de un único signo, acoplamiento bimodal gesto-palabra, acoplamiento unimodal palabra-palabra. Estas habilidades pregramaticales arriban por fin a la articulación gramatical de términos verbales.

Dada la escasa información sobre las habilidades comunicativas bimodales del período de dos palabras, este trabajo pretende aportar evidencia sobre el uso aislado y coordinado de gestos y vocalizaciones durante el mismo desde un estudio evolutivo longitudinal de caso único. El abordaje conjunto de más de una modalidad recupera la idea piagetiana de una función semiótica general,

opuesta a enfoques modularistas donde el lenguaje y la gesticulación serían sistemas de procesamiento inconexos. Entender el desarrollo como *sistema dinámico* supone concebirlo como sujeto a movimiento constante y a una progresión que comprende también estancamientos y retrocesos (progresión no-lineal) que se explican multicausalmente y no desde una relación causal de tipo A entonces B. El enfoque de caso único permite un seguimiento detallado de un sujeto observacional y con ello la posibilidad de atender a las variaciones intra-sujeto, caras a la concepción dinámica por permitir apreciar el cambio en una dimensión que escapa a los estudios transversales.

Método

Participantes

Se observó una díada adulto-niño. Se trata de un niño argentino de la ciudad de Buenos Aires, nacido a término y sin alteraciones ni enfermedades durante el desarrollo. Pertenece a un hogar de clase media, con un hermano tres años mayor, padres profesionales, familia monolingüe. Es el hijo menor del primer autor. La edad del niño en la primera sesión, según la convención piagetiana [año; mes (día)], fue: 1; 7 (13); en la última sesión: 2; 5 (1). El rol del adulto en la díada fue desempeñado alternativamente por la madre, el padre-investigador y una niñera, reflejando de esta manera la realidad de las habilidades comunicativas del niño, libres del sesgo que podrían tener con un interlocutor en particular.

Procedimiento

Se registraron 15 sesiones de interacciones niño-adulto en el hogar del niño y dentro de un contexto natural. Se cuidó en la filmación que el cuerpo de ambos estuviera siempre dentro del cuadro de imagen, especialmente manos, cara y torso, de manera de no perder movimientos o expresiones de ninguno de los dos. Se filmó un total de siete horas y 30 minutos en 15 sesiones de 30 minutos de duración a intervalos de 20 días. Las situaciones de registro tenían temática libre y se desarrollaron dentro de un ambiente lúdico acorde a la edad (juegos de dar y tomar, lectura de cuentos ilustrados, juegos de armar y de insertar, etc.). En todos los casos, el adulto iniciaba los intercambios con propuestas de actividades libradas luego a la espontaneidad del niño. La filmación fue realizada alternativamente por el padre-investigador o por la madre.

En cada sesión videograbada, las intervenciones comunicativas del niño fueron identificadas y contabilizadas. Se consideró comunicación a todo hecho de interacción intencionada y signica (Rivière, 1998), esto es: a las acciones con propósito de contactar al interlocutor mediante la utilización de signos. Se excluyeron los casos de autoemisiones del niño para sí mismo (auto-gesto, verbalizaciones de reconocimiento) por entender que, aunque relevantes para explicar el desarrollo cognitivo, quedan por fuera de nuestra definición de comunicación. Se consideró intervención cada participación comunicativa del

niño en las interacciones con el *partenaire* adulto (gestos, vocalizaciones y composiciones).

Se confeccionó un código de observación conforme al método comparativo constante (Strauss & Corbin, 2002). Se adoptaron primeramente las categorías estándar para el estudio del gesto y de las verbalizaciones durante la infancia temprana. Los gestos se organizaron según la división clásica entre deícticos y representacionales, definidos en la siguiente sección; la expresión oral contempló tanto las vocalizaciones léxicas como las no-léxicas; para las composiciones, se adoptaron las categorías, ya mencionadas, de composición equivalente, complementaria y suplementaria (Volterra et al., 2005), considerando tanto composiciones unimodales (gesto-gesto, vocalización-vocalización) o bimodales (gesto-vocalización). En una segunda fase, con el objetivo de explorar toda la variedad de los recursos comunicativos del niño, se idearon nuevas categorías para las intervenciones cuya clasificación era imposible en las categorías estándar de aquel primer código. Los gestos y las vocalizaciones se clasificaron como *gestos puros* o *vocalizaciones puras* cuando fueron realizados o proferidas, respectivamente, de manera aislada (sin acompañamiento de otro gesto o vocalización), o como *composiciones*, cuando se anexaron a otras unidades (gesto o vocalización, palabra o no-palabra, en número variable). Se consideró composición toda expresión de signos *simultánea* (gesto-vocalización, gesto-gesto [porque manos y cara pueden significar al mismo tiempo]) o *sucesiva* (vocalización-vocalización o gesto-vocalización) que, sin que mediara pausa de más de un segundo, permitiera por contexto comprender su integridad semántica (la copertenencia de los signos como partes de un mensaje). Las composiciones son unimodales cuando sus unidades componentes corresponden solamente a una modalidad (gestual/oral) y bimodales cuando ponen en relación ambas modalidades para una única emisión.

Se aplicó el código de observación a cada una de las sesiones. El contenido de cada sesión fue clasificado independientemente por dos investigadores (los autores de este artículo). Los casos clasificados diferentemente fueron discutidos y argumentados. En las pocas ocasiones en las que no se consiguió acuerdo, la intervención fue descartada. Este procedimiento es consistente con el utilizado en investigaciones de interacción social (Forster & Iacono, 2014; Jonsson et al., 2001).

El código quedó conformado con las siguientes categorías y subcategorías.

Gestos

Los gestos abarcan los tipos: (1) *deíctico* (*pointing* [de señalar], *showing* [de mostrar], *reaching* [de alcanzar] y *request* [de pedir]; el *giving* [gesto de dar], en el límite con la acción, no se ha contabilizado, ya que *hace* y propone al mismo tiempo y por ende no es pura y rigurosamente signo); (2) *representacional* (mimético de la forma total/parcial de un objeto o acción: imitación del vuelo de un avión desplegando los brazos); (3) *conectivo* (de enlace para otros signos: marcación de los dos tiempos de algún episodio — antes/después — mediante un movimiento de la mano a un lado y otro); (4) de *negación/afirmación*

(movimiento giratorio horizontal o sagital con la cabeza); (5) *expresivo* (que dé el tono emocional para entender otro elemento comunicativo: una mueca del rostro, un golpe sobre alguna superficie subrayando una expresión verbal); y (6) *arbitrario* (mano agitada para comunicar ‘chau’). Los primeros dos tipos y el último corresponden a una categorización tradicional (por ejemplo, Iverson & Thal, 1998); el tercero a McNeill (1992); el cuarto y el quinto son aportes clasificatorios de esta investigación.

Vocalizaciones

Se computaron como vocalizaciones tanto las palabras como las expresiones orales por fuera del léxico (no-palabras) en las que, no obstante, se reconocieron intenciones comunicativas y aptitud para integrarse en compuestos de significación. Las no-palabras — interjecciones expresivas (‘¡eh!’, ‘¡oh!’, ‘¡ay!’) y modulaciones no-expresivas (e.g., ‘ta’, o bien ‘ta-ta-ta’) emitidas efectivamente para interactuar — han sido revalorizadas en los estudios de adquisición del lenguaje (Karousou & López Ornat, 2013), reconociéndoseles un valor predictivo inclusive mayor al de los gestos en las futuras habilidades lingüísticas. Con las modulaciones no-expresivas, que podían ser monosilábicas o multisilábicas, se procedió contabilizando las primeras como unidades sin más, mientras que a las segundas se las dividió en unidades silábicas simples, salvo cuando presentaban una forma constante (unidad trisilábica *ta-ta-ta* temporalmente separada — por más de un segundo — de otras emisiones) y eran empleadas en por lo menos dos sesiones dentro de contextos semánticamente definidos (cumpliendo así con la definición estandarizada de protopalabras).

Composiciones

Las composiciones se clasificaron en función de (a) tipo y cantidad de componentes, y (b) la colaboración semántica entre los componentes. La clasificación (a) distingue las composiciones: gesto-gesto (G-G), gesto-vocalización (G-V), vocalización-vocalización (2V), gesto-gesto-vocalización (2G-V), gesto-vocalización-vocalización (G-2V), gesto-gesto-vocalización-vocalización (2G-2V) y composiciones unimodales de tres, cuatro ó más vocalizaciones. La clasificación (b) distingue las composiciones *equivalente* (unidades semánticamente idénticas), *complementaria* (semánticamente convergentes), *suplementaria* (semánticamente diferentes), *emotiva* (donde ora la gesticulación, ora la vocalización aportan simplemente un énfasis sobre el significado del otro constituyente y no, como en las anteriores, contenido informativo — ejemplo: el nombre de un juguete acompañado por gesto expresivo de sorpresa ante su hallazgo) y *dudosa* (donde no pudo decidirse a cuál de las otras cuatro categorías convenía asignar una composición). Las composiciones emotivas no están contempladas en los estudios de este tipo y han sido incorporadas en una segunda fase del proceso de codificación al haber reconocido su importancia.

Las vocalizaciones no-palabra reduplicativas del tipo ‘ta-ta-ta’ no se contabilizaron, como parte de las composiciones por considerarlas simplemente un ejercicio vocálico (a veces muy largo), aunque sí fueron contabilizadas, como se dijo, entre las emisiones orales. Sólo las interjecciones y las protopalabras fueron consideradas vocalizaciones integrantes de composiciones.

Las composiciones están delimitadas por su unidad semántica y por su relativa coincidencia temporal, como se ha explicado. En cuanto a lo primero, los signos intervinientes en la composición debían tener la misma referencia (*pointing* + verbalización ‘acá’ = G-V), o vincular dos referentes (‘yo’ + ‘pelota’ = 2V), o predicar un atributo de un referente (*pointing* + ‘esto’ + ‘grande’ = G-2V) o expresar una emoción respecto de un referente (‘no está’ + gesto de sorpresa = G-2V). La pauta temporal supone coincidencia pero no perfecta sincronía (un criterio imposible si quiere abarcarse las composiciones de dos o más vocalizaciones, las que implican necesariamente una disposición sucesiva de los elementos sígnicos, ya que pertenecen a una misma modalidad). Se consideraron pautas límite de una composición sígnica el cambio absoluto de referencialidad semántica y un desfase temporal mayor a un segundo. El orden de los elementos de la composición no se consideró relevante dado que en el período estudiado ellas se entienden como yuxtaposiciones sin evidencia clara de gramaticalidad (López Ornat, 1999). Así pues, las composiciones no se consideraron pares ordenados (o triplos, etc.) sino simplemente acoplamientos sígnicos.

Resultados

Gestos y vocalizaciones

La [Tabla 1](#) muestra la frecuencia de los gestos y las vocalizaciones en sus diferentes tipos de ocurrencia (pura y en composición) contabilizadas durante el período observado. A lo largo de todo el período estudiado se registró un neto predominio de las intervenciones de elementos puros (gestos o vocalizaciones aisladas) por sobre las de composición, con una diferencia del 56.3% en la sesión 1 y del 20.6% en la sesión 15 (en promedio una diferencia del 33.8%). Las composiciones (gesto-gesto, gesto-vocalización, vocalización-vocalización) nunca llegan a superar las intervenciones de un solo elemento.

La [Tabla 2](#) muestra que la frecuencia de vocalizaciones (puras, en composición con gesto o vocalización y totales) es superior a la de gestos (puros, en composición con vocalizaciones o con otros gestos y totales) a lo largo de todo el período. Esto señala una preferencia general por la modalidad oral en las intervenciones de cada una de las sesiones. Se identificó una correlación positiva entre los gestos totales y las vocalizaciones totales ($r_s = .554$; $p < .05$), lo que estaría mostrando que los comportamientos de una y otra modalidad, más allá del predominio de la expresión oral, se mantienen en proporción y que el paulatino aumento de los signos orales no se realiza a expensas de los signos gestuales.

Para la modalidad gestual, los gestos integrados en composiciones constituyen 85.2% del total de gestos, con una diferencia en el promedio de frecuencias a lo largo de las 15 sesiones del 62.2 por sobre el promedio en la frecuencia de los

Tabla 1. Frecuencia de intervenciones con elementos puros (gestos o vocalizaciones) y composiciones.

Sesión	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Totales
Intervenciones Totales	403	238	238	313	339	245	311	421	358	229	350	465	455	432	310	5,105
Intervenciones con Elementos Puros	315	175	137	214	192	163	226	264	241	166	248	308	298	286	187	3,420
Intervenciones con Elementos Puros %	78.1	73.5	57.6	68.3	56.6	66.5	72.7	62.7	67.3	72.5	70.9	66.2	65.5	66.2	60.3	66.9
Intervenciones con Composiciones	88	63	101	99	147	82	85	157	117	63	102	157	157	146	123	1687
Intervenciones con Composiciones %	21.8	26.5	42.4	31.6	43.4	33.5	27.3	37.3	32.7	27.5	29.1	33.8	34.5	33.8	39.7	33.0

Tabla 2. Frecuencia de gestos y vocalizaciones puras o integradas a composiciones.

Sesión	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Totales
Gestos Totales	68	62	99	98	143	57	59	114	81	73	92	125	111	121	87	1,390
Gestos Puros	7	8	2	13	14	6	0	12	18	27	18	16	26	21	17	205
Gestos Composiciones (gesto-gesto, gesto-voc.)	61	54	97	85	129	51	59	102	63	46	74	109	85	100	70	1,185
% Gestos Composiciones /Gestos Totales	89.7	87.1	98	86.8	92.3	89.5	100	89.5	77.8	63.1	80.5	87.2	76.6	82.7	90	85.2
Vocaliz. Totales	455	245	247	326	367	295	353	491	421	234	380	538	537	527	416	5,832
Vocalizaciones Puras	308	167	135	201	178	157	226	252	223	139	230	292	272	265	170	3,215
Vocalizaciones Compos. (gesto-voc., voc.-voc.)	147	78	112	125	189	138	127	239	198	95	150	246	265	262	246	2,617
% Vocaliz. Composiciones /Voc. Totales	32.3	31.8	45.3	38.3	51.4	46.8	35.9	48.6	47	40.6	39.5	45.7	49.3	49.7	59.1	44.8

gestos puros. Se aprecia un leve incremento de los gestos puros hacia la segunda mitad del período estudiado, que a partir de la sesión 8 no descienden de los dos dígitos y en tres ocasiones tienen frecuencias superiores a 20 (sesiones 10, 13 y 14).

Para las vocalizaciones, y al contrario de lo que sucede con los gestos, se observa un predominio de las intervenciones puras por sobre las de vocalizaciones integradas en composiciones. Esta relación sólo se invierte en la última sesión, donde las vocalizaciones en composición, luego de obtener frecuencias muy similares con las vocalizaciones puras durante las sesiones previas (12, 13 y 14), logran mayor frecuencia en proporción de 246 a 170.

La comparación entre gestos puros y vocalizaciones puras arroja un resultado abrumadoramente superior para la modalidad oral: durante las 15 sesiones de observación, el promedio de la frecuencia de vocalizaciones puras es superior al promedio de la frecuencia de gestos puros por 200.7 (ver [Tabla 2](#)). La comparación entre gestos integrados en composiciones y vocalizaciones integradas en composiciones revela también una diferencia en el promedio de frecuencias de 95.5 favorable a estas últimas. Hacia la segunda mitad del período la diferencia de valores se agudiza y las vocalizaciones-composición superan las 100 en la mayoría de estas sesiones (excepto en las sesiones 10 y 11).

Composiciones

Composiciones de acuerdo con el tipo y la cantidad de componentes

La [Tabla 3](#) muestra las composiciones encontradas según tipo (gesto o vocalización) y cantidad de componentes en cada una. A lo largo de todo el período, el tipo de composición predominante es el de gesto-vocalización (G-V = 53.9%). Sumando las frecuencias de todos los tipos de emisiones bimodales (G-V, 2G-V, 2V-G, 2G-2V y 3V-G), el resultado es superior al de la frecuencia total de casos de composición unimodal (2V, 3V y 4V, o G-G, estas de $n = 4$). Se obtuvieron 1,152 composiciones bimodales (68.3% de las composiciones totales del sujeto) contra 531 composiciones con solo vocalizaciones (31.5% del total de

Tabla 3. Frecuencia de composiciones de acuerdo con el tipo y la cantidad de componentes (G/gesto; V/vocalización).

Sesión	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Totales	%
G-V	44	50	90	77	112	40	47	77	44	35	59	81	69	55	30	910	53.9
G-G	0	0	1	0	0	0	0	1	1	0	0	0	0	0	1	4	0.2
2G-V	1	0	0	1	1	0	0	2	0	1	1	2	0	3	1	13	0.8
2V	16	8	3	11	13	18	23	54	47	15	24	40	58	42	35	407	24.1
2V-G	11	2	5	4	13	10	11	18	15	7	11	18	11	17	26	179	10.6
2G-2V	0	1	0	1	0	0	0	0	0	1	1	2	1	3	2	12	0.7
3V	12	2	2	5	6	13	3	3	8	3	6	11	13	8	17	112	6.6
3V-G	4	0	0	0	2	1	1	1	2	0	0	2	3	16	6	38	2.2
4V	0	0	0	0	0	0	0	1	0	1	0	1	2	2	5	12	0.7

Tabla 4. Frecuencia de composiciones totales gesto-palabra y palabra-palabra en comparación con frecuencia de composiciones totales de gesto-vocalización y vocalización-vocalización. (G/gesto; V/vocalización; P/palabra)¹.

Sesión	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Totales
G-V	60	53	95	83	128	51	59	98	61	44	72	105	84	94	65	1,152
G-P	51	41	93	69	91	46	50	91	59	34	69	92	63	82	59	990
V-V	28	10	5	16	19	31	26	58	55	19	30	59	73	52	57	529
P-P	0	0	0	5	5	4	10	20	38	7	16	32	47	34	42	260

Nota: ¹La tabla contabiliza, a diferencia de la [Tabla 3](#), la totalidad de composiciones G-V Y V-V, sin atender a la cantidad de componentes de una u otra modalidad. De esta manera, G-V comprende aquí G-V, 2G-V, 2V-G, 2G-2V Y 3G-V de la [Tabla 3](#), mientras que V_V comprende 2V, 3V y 4V de la [Tabla 3](#).

composiciones). El 0.23% restante corresponde a casos de composición sin vocalización (G-G).

Se han identificado las siguientes correlaciones, calculando el coeficiente de asociación de Spearman: (a) unimodal, entre 2V y 3V ($r_s = .523$; $p < .05$); y (b) bimodales, entre 2V-G y 3V-G ($r_s = .729$; $p < .01$) y entre 3V y 3V-G ($r_s = .747$; $p < .01$). En conjunto, estos resultados indicarían que el patrón de crecimiento en la cantidad de componentes no ocurre solamente en la modalidad oral sino que las composiciones bimodales, con intervención del gesto, lejos de volverse recesivas, acompañan el aumento de las comunicaciones del niño.

Según se indica en el Código de Observación, la denominación de vocalización abarca tanto palabras como no-palabras. Debido a que las palabras contribuyen decisivamente al proceso de gramaticalización, dentro de las composiciones gesto-vocalización se ha distinguido, como subconjunto, la categoría gesto-palabra, y dentro de las composiciones vocalización-vocalización se ha distinguido, como subconjunto, la categoría palabra-palabra. La distinción es importante para contabilizar las unidades léxicas integradas a composiciones y seguir de cerca estas composiciones en su marcha hacia la *unimodalización* (verbal) sobre la que se desarrolla la gramática. Las frecuencias de estas dos variables se detallan en la [Tabla 4](#). Conforme al curso esperable de la evolución lingüística del niño, las composiciones palabra-palabra aumentan con el correr del tiempo, pero no a expensas de las composiciones gesto-palabra.

Composiciones de acuerdo con el tipo de colaboración entre los componentes

Se evaluó si los componentes de gesto y vocalización se integraban en composiciones equivalentes (CEQ), complementarias (CC), suplementarias (CS), emotivas (CE) y dudosas (CD). La [Tabla 5](#) muestra las frecuencias de cada tipo identificado.

Se observó una frecuencia mayor de composiciones equivalentes, las cuales fueron más pronunciadas en la segunda mitad de las sesiones. En paralelo, si bien las composiciones suplementarias se encuentran presentes desde el comienzo, no muestran un incremento progresivo. Un resultado de interés lo constituye la alta

Tabla 5. Frecuencia de composiciones de acuerdo con el tipo de colaboración entre los componentes.

Sesión	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Totales	%
CEQ	48	60	27	61	39	49	49	110	91	49	77	78	74	81	63	956	56.6
CC	8	22	14	41	22	21	15	11	8	6	1	10	30	11	21	241	14.3
CS	9	7	5	20	9	0	6	8	8	0	14	24	22	17	11	160	9.5
CE	20	5	8	9	20	12	6	22	6	1	5	39	17	26	15	211	12.5
CD	3	7	9	16	9	0	9	6	4	7	5	6	14	11	13	119	7.0

frecuencia relativa de las composiciones emotivas (12.5%), consideradas por primera vez en este estudio. Su frecuencia es mayor que la de las composiciones suplementarias (9.5%), que de acuerdo con investigaciones previas debían ser predominantes, y apenas menor a la de composiciones complementarias (14.3%).

Discusión

Gestos y vocalizaciones

Durante todo el período estudiado pudo observarse una mayor frecuencia de vocalizaciones puras que de gestos y composiciones (ver [Tabla 2](#)). Ello indica una preferencia constante por la modalidad oral, lo cual es coherente con el hecho de que esta modalidad es también la primera opción con la que el niño expresa sus estados de necesidad y de emotividad. No obstante, la correlación entre gestos totales y vocalizaciones totales indica que el comportamiento de las dos modalidades comunicativas es semejante con independencia de las diversas frecuencias. El aumento de la comunicatividad oral del niño en el curso de una sesión determinada se acompaña del aumento de la comunicatividad gestual. Es como si, en su predisposición a interactuar, el niño utilizara ambos recursos, con una marcada preferencia por la oralidad pero sin descuidar la alternativa de los gestos y existiendo entre ambas una proporción constante.

Pese a haber estudiado el período de dos palabras, son sin embargo las vocalizaciones puras las expresiones dominantes. Ello supone que la denominación ‘de dos palabras’ sólo indica la *ocurrencia* de pares de elementos verbales sin aludir a su frecuencia. No obstante, en la segunda mitad del período se aprecia un incremento importante de las composiciones palabra-palabra (ver [Tabla 4](#)). Ello revela que, sobre un fondo de vocalizaciones invariablemente dominantes, las composiciones de unidades de palabra, necesarias para la comunicación verbal gramatical, aumentan junto con la variable tiempo como una especialización paulatina dentro del conjunto de las posibilidades de expresión oral.

La relación inversa en el comportamiento entre composiciones palabra-palabra y elementos puros de la oralidad sugiere el acomodamiento de los diferentes recursos expresivos al desarrollo normal del sujeto, en cuyo horizonte se encuentra el lenguaje articulado, para el cual es necesario el establecimiento de composiciones orales sobre las que puedan operar las reglas de articulación gramatical.

Los gestos puros, elementos poco presentes al comienzo, se incrementan en las últimas sesiones. Ello puede deberse a un aumento general de la comunicatividad en el niño, de acuerdo con la cual, y sobre una explicación bimodal, estaría justificado el crecimiento de frecuencias conjunto para gesto y oralidad. Esta interpretación se condice con la correlación entre gestos totales y vocalizaciones totales comentada más arriba.

Composiciones

Se verificó un neto predominio del tipo gesto-vocalización, un resultado interesante ya que pone de relieve que, incluso en el período de dos palabras, las composiciones están dominadas por la bimodalidad. Así, la predominancia de los elementos de expresión oral dentro de la totalidad de las intervenciones del niño no se acompaña de un comportamiento igualmente predominante de composiciones exclusivamente orales. Ello señala que la oralidad se usa en la mayoría de intervenciones o bien sola o bien unida a gesto, pero no ligada — todavía — a otro elemento oral.

El crecimiento de las composiciones de modalidad exclusivamente oral a partir de la mitad del período (ver [Tabla 2](#)) es coherente con el progreso del lenguaje. Si se considera al mismo tiempo que las composiciones de gesto-vocalización son dominantes durante todo el período, la transformación cognitiva encaminada hacia la oralidad gramatical podría explicarse, a este nivel, como un proceso dinámico de incorporación gradual de los elementos de una modalidad en una estructura preexistente de procesamiento bimodal. Este proceso es dinámico en cuanto el crecimiento de casos unimodales orales no reemplaza la primera forma de composición de dos elementos, más bien constituye una adquisición no-sustitutiva. De hecho, la palabra siempre tendrá soporte en esta estructura bimodal, y en el gesto un auxiliar comunicativo permanente (McNeill, 1992). La conexión y mutua asistencia entre la expresión gestual y la oral apuntan en favor de un dispositivo de procesamiento permeable a ambos tipos de unidad comunicativa y no encapsulado, como pretende la versión modularista del lenguaje.

El crecimiento de las composiciones palabra-palabra, sostenido pero irregular, no supera a las composiciones de gesto-palabra (ver [Tabla 4](#)) ni resiente su frecuencia relativamente estable. Ello señala que el aumento de composiciones de naturaleza unimodal, el destino final de la expresión llamada a ser lingüística, sucede simplemente porque la modalidad oral es una opción privilegiada en las interacciones comunicativas de la especie humana.

Frecuencias de composiciones de acuerdo con el tipo y cantidad de componentes

Desde el inicio del período se observan composiciones de gesto-palabra, pero no de palabra-palabra. El primer ejemplo de este último tipo se registra recién en la sesión 4 ($n = 5$), y de allí crece de manera irregular pero constante. Las composiciones gesto-gesto apenas suman cuatro en todo el período, lo que lleva a pensar que la expresión gestual se encuentra especialmente cómoda en formato puro o en

composición con vocalizaciones (cuestión que deja ver hasta qué punto la modalidad oral es un pivote para las expresiones complejas).

Hay, asimismo, desde el comienzo del registro, intervenciones de modalidad oral de más de dos vocalizaciones (3V y 3V-G — considerando vocalizaciones de todo tipo, palabra y no-palabra). Este dato, contrastado con la ausencia de composiciones palabra-palabra, muestra que en principio la oralidad sirve a la expresión de elementos no-léxicos con elementos léxicos, y que sólo más tarde canaliza la comunicación verbal stricto sensu.

Las composiciones de dos elementos unimodales y bimodales suman el 78.3% del total de las composiciones (ver [Tabla 3](#)), lo que marca con claridad que las habilidades comunicativas del período se encuentran cómodas con el patrón intermodal de dos *slots* o espacios de llenado. La capacidad de acoplar dos elementos se consolida durante el período y se proyecta en paralelo como plataforma para otras composiciones de más elementos tanto unimodales como bimodales.

Si se considera, por un lado, las correlaciones entre distintas composiciones con un número igual o cercano de elementos componentes (2V/3V, 2V-G/3V-G, 3V/3V-G), y, por otro, el hecho de que desde la sesión 12 se produce un incipiente crecimiento de las composiciones de 2V-2G, 3 V-G y 4V, ello en conjunto sugiere que el incremento ocurre en la cantidad de los componentes (cuatro) con independencia de la modalidad. Todo esto abona la idea de que la comunicación del niño se apoya indistintamente en todos los recursos (canales) a su disposición y que la autonomización relativa de la palabra dentro del sistema gramatical de una lengua es el producto de una previa colaboración intermodal. Aunque este punto deberá ponerse entre paréntesis por la frecuencia escasa de algunas de estas composiciones (sólo 12, para las variables de cuatro elementos, en todo el período estudiado), las intervenciones del niño verifican en efecto en general un parejo crecimiento en el número de sus elementos sin importar la modalidad a la que pertenecen.

Si tomamos en cuenta que el conjunto de composiciones de palabra describe una tendencia ascendente, el progreso general de las habilidades comunicativas del niño observado parece conforme con la expectativa general del desarrollo del lenguaje, que supone una ampliación progresiva del vocabulario y las composiciones entre sus términos.

Frecuencias de composiciones de acuerdo con el tipo de colaboración entre componentes

En relación con el tipo de colaboración semántica entre los componentes de una comunicación, las composiciones de mayor frecuencia son las equivalentes, donde ambos elementos (aquí gesto y palabra, inevitablemente) refieren lo mismo con la misma carga de significado (deíctico pronominal y gestual: *pointing* + ‘esto’). Su predominio se extiende a lo largo de todo el período con un total del 56.6% de las intervenciones de esta clase. Ello es de destacar porque en estudios precedentes (Volterra et al., 2005) se informa que a finales de la etapa de holofrases las

composiciones de tipo suplementario han superado a las complementarias y equivalentes. Si esos resultados describían una razonable progresión de menor a mayor dificultad en el procesamiento de la información, pasando de un primer formato de composiciones equivalentes a una segunda variedad de composiciones complementarias (un referente aludido diversamente a través de dos unidades de significado diferente y de modalidades diferentes, gestualidad y oralidad) y luego a una composición suplementaria que integra dos contenidos semánticamente autónomos (dos referentes que cumplen *avant la lettre* las funciones de sujeto y predicado), los resultados de nuestra investigación se alejan de tal secuencia 'lógica'.

Sucede también que las composiciones complementarias decrecen sin haber sobrepasado nunca la frecuencia de las suplementarias (a las cuales deberían más tarde ceder el lugar predominante — Volterra et al., 2005). Las composiciones suplementarias, por su parte, están presentes desde el comienzo, pero no se aprecia en ellas un aumento de tendencia sostenida y al final de las sesiones su frecuencia es todavía inferior a la de las composiciones equivalentes. Estos datos concretos vuelven a desafiar la transición entre composiciones equivalentes — complementarias — suplementarias. Los resultados obtenidos en un estudio de caso como el nuestro podrían sugerir que la preponderancia de las composiciones equivalentes es tan sólo una cuestión de estilo comunicativo, pero no cabe descartar que aquella secuencia merezca una consideración más detallada y nuevas investigaciones.

La falta de un rol predominante para las composiciones de tipo suplementario es doblemente inesperada si se estima que, de una manera novedosa, en la investigación hemos optado por incluir no solamente los acoplamientos gesto-vocalización, sino también aquellos exclusivamente orales, no contemplados en otros estudios precedentes. Esta variante invitaba a pensar que el predominio de composiciones de tipo suplementario sería todavía mayor, dado que a la superior frecuencia de composiciones suplementarias bimodales de gesto-palabra se adicionarían composiciones palabra-palabra. Sin embargo, esta definición más comprensiva de las composiciones suplementarias no redundó en una mayor frecuencia.

La frecuencia de las composiciones emotivas, consideradas por primera vez en este estudio, en relación con las demás (12.5%) induce a pensar que sería conveniente, en adelante, que los estudios de gesto y primeras composiciones de gesto-palabra las incorporasen. Posiblemente este descuido en los estudios previos tenga que ver con que se encuentran, en su mayor parte, conformados por los intereses de la adquisición lingüística, con el colateral sesgo excluyente de las manifestaciones emotivas no-referenciales.

La persistencia, a lo largo de todo el período, de las composiciones emotivas, que incluyen invariablemente algún elemento de expresión emocional (interjección, gesto emotivo), junto con la de las composiciones equivalentes, enfrentadas al crecimiento de las composiciones suplementarias, que implican acoplamientos cuasi-proposicionales, sugiere que los usos de la oralidad más simples y originarios (dicho por las composiciones emotivas, que manifiestan

estados de ánimo) y los usos de las composiciones más elementales conviven con el creciente número de composiciones que requieren el manejo de operaciones de mayor complejidad cognitiva (como es la predicación). Así, en el proceso de sofisticación de las herramientas comunicativas, lo más primitivo convive con las nuevas estrategias (tal como puede constatarse en los adultos, para quienes el dominio léxico no anula las interjecciones). Esta perduración de los primeros usos comunicativos de la oralidad junto con otros ulteriores nos permite encuadrar el desarrollo de las habilidades semióticas como un proceso dinámico de progreso y conservación, donde lo nuevo está en constante enlace con lo anterior y ambas estrategias cognitivas se interpenetran (para el caso, el léxico expresa emociones y las interjecciones se acogen a estructuras lingüísticas formalmente correctas).

Consideraciones generales

La investigación ha encontrado una preferencia por la modalidad oral frente a la gestual, para aquellas intervenciones puras o de único componente, y una preferencia por las composiciones bimodales respecto de las unimodales. Si la bimodalidad es la norma durante el período holofrástico, la novedad de nuestro estudio es que ello parecería continuar en el período de dos palabras. El pivote que la construcción gesto-palabra ofrece a la intención de interactuar del niño no se volvería recesivo con el incremento de composiciones exclusivamente de palabra, las cuales se desarrollarían como una segunda ruta en paralelo. La bimodalidad presente en las comunicaciones del período, lejos de ser incidental, abonaría la tesis de una estructura cognitivo-comunicativa como la propuesta por McNeill (1992).

En cuanto a las limitaciones de nuestra investigación es importante destacar que, como estudio de caso único, no permite desde luego generalizaciones ni explicaciones de fondo, y que la evidencia del estudio longitudinal no está libre del riesgo de reflejar sesgos de estilo personal del sujeto, diferencias madurativas individuales y la incidencia no controlada de variables externas. Con todo, los resultados de un trabajo de este tipo son, en la temática que nos ocupa, ‘imprescindibles’ (López Ornat, 1994, p. 8), porque reflejan de manera pormenorizada el proceso de desarrollo del fenómeno estudiado, atendiendo a los progresos tanto como a los puntos de detención e incluso los de retroceso. Futuros estudios deberán determinar si la desorganización de la secuencia de composiciones equivalentes — complementarias — suplementarias ha sido un hecho idiosincrático del estudio de caso o bien responde a cierta regularidad. Otra limitación del estudio es soslayar una descripción cualitativa de los procesos referidos. Se encuentra en preparación otro trabajo que suministra este tipo de datos. Finalmente, la presente investigación, concentrada en las composiciones y su vínculo con las llamadas formas de ocurrencia puras, tampoco informa acerca de las frecuencias de los tipos de gesto y de palabra que se han presentado en el período. Esa información es material para otro artículo relacionado con el presente.

El abordaje multimodal de la primera infancia y en particular del desarrollo ontogenético de las habilidades comunicativas de los niños en proceso de adquisición del lenguaje revela que los factores intervinientes son de más amplio espectro que lo que se manifiesta en los enfoques limitados a los progresos verbales. Redimensionar multimodalmente este proceso supone no sólo apreciar desde un distinto mirador los usos espontáneos que los niños realizan de su dotación expresiva natural, sino también la posibilidad de proveer de herramientas polifacéticas a los profesionales del campo educacional, cuyas estrategias de trabajo se verán enriquecidas con el aporte de la nueva perspectiva. Los programas de intervención clínica y psicopedagógica temprana deberían reexaminar sus lineamientos y reasignar al gesto y a las composiciones de gesto-palabra un espacio mayor, considerándolos como un soporte central en la adquisición del lenguaje, y a ésta a su vez como un proceso ni lineal, ni modular, ni unimodal, sino, al contrario, como un desarrollo sujeto a aceleraciones y detenciones y abierto a distintas representaciones simbólicas cualitativamente diversas (por su también diferente vehículo expresivo o variedad modal, gesto o vocalización). Creemos, en definitiva, que el presente estudio — como alternativa a pensar el lenguaje en calidad de dispositivo autónomo impermeable a otros influjos cognitivos — permite concebir el desarrollo de las habilidades semióticas del niño como un sistema dinámico de pleno derecho.

Acknowledgements / Agradecimientos

This study was performed as part of project *D15S05 *Pragmatics and semantics in the early development of human communication. The intermodal relationships of speech and gesture during first verbal compositions* funded by the Universidad Argentina de la Empresa, and from project ***(ANPCyT) PICT–2013-0368 *Communicative musicality in temporal arts and early childhood.* /Esta investigación se ha realizado dentro del proyecto *D15S05 Pragmática y semántica en el desarrollo temprano de la comunicación humana. Las relaciones intermodales de habla y gesto durante el período de primeras composiciones verbales *subvencionado por la Universidad Argentina de la Empresa, y del proyecto ***(ANPCyT) PICT–2013-0368 Musicalidad comunicativa en las artes temporales y la infancia temprana.*

Disclosure statement

No potential conflict of interest was reported by the authors. / *Los autores no han referido ningún potencial conflicto de interés en relación con este artículo.*

References / Referencias

- Butcher, C., & Goldin-Meadow, S. (2000). Gesture and the transition from one- to two-word speech: When hand and mouth come together. In D. McNeill (Ed.), *Language and gesture* (pp. 235–258). New York, NY: Cambridge University Press.
- Capirci, O., & Volterra, V. (2008). Gesture and speech: The emergence and development of a strong changing partnership. *Gesture*, 8, 22–44. doi:10.1075/gest.8.1.04cap
- Chomsky, N. (1984). *Modular approaches to the study of the mind*. San Diego, CA: San Diego State University Press.
- Fodor, J. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.

- Forster, S., & Iacono, T. (2014). The nature of affect attunement used by disability support workers interacting with adults with profound intellectual and multiple disabilities. *Journal of Intellectual Disability Research*, 58, 1105–1120. doi:10.1111/jir.2014.58.issue-12
- Iverson, J., Capirci, O., & Caselli, M. (1994). From communication to language in two modalities. *Cognitive Development*, 9, 23–43. doi:10.1016/0885-2014(94)90018-3
- Iverson, J., & Thal, D. (1998). Communicative transitions: There's more to the hand than meets the eye. In A. Wetherby, S. Warren, & J. Reichle (Eds.), *Transitions in prelinguistic communications* (pp. 59–86). Baltimore, MD: Paul H. Brookes.
- Jonsson, C.-O., Clinton, D., Fahrman, M., Mazzaglia, G., Novak, S., & Söhrus, K. (2001). How do mothers signal shared feeling-states to their infants? An investigation of affect attunement and imitation during the first year of life. *Scandinavian Journal of Psychology*, 42, 377–381. doi:10.1111/1467-9450.00249
- Karmiloff-Smith, A. (1995). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.
- Karousou, A., & López Ornat, S. (2013). Prespeech vocalizations and the emergence of speech: A study of 1005 Spanish children. *The Spanish Journal of Psychology*, 16, 1–21. doi:10.1017/sjp.2013.27
- López Ornat, S. (1994). La metodología de la investigación longitudinal. In S. López Ornat, A. Fernández, P. Gallo, & S. Mariscal (Eds.), *La adquisición de la lengua española* (pp. 3–12). Madrid: Siglo XXI.
- López Ornat, S. (1999). La adquisición del lenguaje. Nuevas perspectivas. In F. Cuetos, & M. De Vega (Eds.), *Psicolingüística del Español* (pp. 469–533). Madrid: Trotta.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.
- Rivière, Á. (1998). Tratamiento y definición del espectro autista I: Relaciones sociales y comunicación. In Á. Rivière, & J. Martos (Eds.), *El tratamiento del autismo. Nuevas perspectivas* (pp. 61–106). Madrid: IMSERSO.
- Serra, M., Serrat, E., Solé, R., Bel, A., & Aparici, M. (2000). *La adquisición del lenguaje*. Barcelona: Ariel.
- Strauss, A., & Corbin, J. (2002). *Bases de la investigación cualitativa. Técnicas y procedimientos para desarrollar la teoría fundamentada*. Medellín: Editorial Universidad de Antioquia.
- Volterra, V., Caselli, M. C., Capirci, O., & Pizzuto, E. (2005). Gesture and the emergence and development of language. In M. Tomasello, & D. Slobin (Eds.), *Beyond nature-nurture: Essays in honor of Elizabeth Bates* (pp. 3–40). Hillsdale, NJ: Lawrence Erlbaum Associates.