



LUND UNIVERSITY

Developmental perspectives on the expression of motion in speech and gesture. A comparison of French and English

Hickmann, Maya; Hendriks, Henriette; Gullberg, Marianne

Published in:
Language, Interaction and Acquisition

DOI:
[10.1075/lia.2.1.06hic](https://doi.org/10.1075/lia.2.1.06hic)

2011

[Link to publication](#)

Citation for published version (APA):
Hickmann, M., Hendriks, H., & Gullberg, M. (2011). Developmental perspectives on the expression of motion in speech and gesture. A comparison of French and English. *Language, Interaction and Acquisition*, 2(1), 129-156.
<https://doi.org/10.1075/lia.2.1.06hic>

Total number of authors:
3

General rights

Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Developmental perspectives on the expression of motion in speech and gesture: A comparison of French and English

Maya Hickmann, Henriëtte Hendriks & Marianne Gullberg

Abstract

Recent research shows that adult speakers of verb- vs. satellite-framed languages (Talmy, 2000) express motion events in language-specific ways in speech (Slobin 1996, 2004) and co-verbal gestures (Duncan 2005; Kita & Özyurek 2003; McNeill 1992). Although such findings suggest cross-linguistic differences in the expression of events, little is still known about their implications for first language acquisition. This paper examines how French and English adults and children (ages four and six) express Path and Manner in speech and gesture when describing voluntary motion presented in animated cartoons. The results show that English adults conflate Manner+Path in speech more often than French adults who frequently talk about Path only. Both groups gesture mainly about Path only, but English adults also conflate Manner+Path into single gestures, whereas French adults never do so. Children in both languages are predominantly adult-like in speech and gesture from age four on, but also display developmental progressions with increasing age. Finally, speech and gestures are predominantly co-expressive in both language groups and at all ages. When modalities differ, English adults typically provide less information in gesture (Path) than in speech (Manner+Path; ‘Manner modulation’ phenomenon), whereas French adults express complementary information in speech (Manner) and gesture (Path). The discussion highlights theoretical implications of such bi-modal analyses for acquisition and gesture studies

Key words: multimodality; speech-gesture co-expressivity; typology and acquisition; space; manner of motion; path of motion

1. Introduction

Many studies in the past few decades have examined the expression of motion across languages, as well as co-verbal gestures, specifically during speech about motion. A number of hypotheses have been proposed regarding the role of these gestures: first, they might be precursors to speech and / or facilitators for expression; second, they might add information not expressed in speech; and third, they might mirror the information expressed in speech. Testing these hypotheses requires a cross-linguistic perspective, given that linguistic systems seem to highlight different components of motion events, thereby most likely also influencing the information that is expressed in gesture. Finally, from a

developmental perspective, an important question concerns whether children and adults speaking the same language use speech and gesture in the same way or if they assign different functions to the two modes of expression. These questions are addressed in the present paper, with particular attention to French and English.

1.1. Speaking about motion across languages

Talmy (2000) has explored the linguistic means whereby basic semantic components of motion (Motion itself and Path) as well as additional information (Cause and Manner) are typically expressed across languages. He distinguishes at least two classes of languages, illustrated in English (1) and French (2) with motion events implying changes of location:¹ *satellite-framed* languages that typically express Path in satellites (e.g. Germanic) and *verb-framed* languages that typically express Path in the verb (e.g. Romance). Note that the locus of Path information affects the expression of Manner. When Path is marked in satellites (English (1)), the verb is free to express Manner. However, when the main verb encodes Path (French (2)), other peripheral means must be used for Manner, which is relegated to the sentence periphery (e.g. in a gerund) or frequently not expressed at all.

- (1) *Oscar runs* [Manner] *into* [Path] *the kitchen*.
- (2) *Oscar entre* [Path] *dans la cuisine* (*en courant* [Manner]).
 ‘Oscar enters in[to] the kitchen ([by] running)’

Thus, languages seem to show preferences with regard to what information is more readily expressed. These cross-linguistic differences are also reflected in speech-associated gestures. Many studies show that adult speakers of different languages talk and gesture differently about Path and Manner depending on whether they speak a verb- or satellite-framed language (e.g. Duncan 2005; Kita & Özyürek 2003; McNeill 1992). For example, in Turkish (verb-framed) speakers typically express Path and Manner in separate clauses (e.g. *while running*) and more often perform separate gestures for Path (direction) and Manner (e.g. wiggling fingers for running legs). In comparison English speakers typically produce Manner and Path in a single clause (*runs out*) and sometimes in one gesture, thereby conflating Manner and Path.

1.2. Implications for language and cognitive development

¹ This typological distinction is most relevant when motion events imply changes of location, whether gradual (e.g. upward/downward motion) or categorical (e.g. boundary crossing), rather than when they take place within a general location (e.g. *Il court dans le jardin* / *He is running in the garden*).

Research has now begun to examine how children discover adult preferences on their way to becoming native speakers. If children start out referring to specific types of information on the basis of prelinguistic knowledge, one should find similarities at early ages across all languages, followed by a later process of tuning in to the target language. Earlier studies (Johnston & Slobin 1979) as well as more recent infant research (Hespos & Spelke 2004, 2007; Mandler 1998; Spelke 2003) indeed suggest that this might be the case. However, other studies (Allen et al. 2007; Berman & Slobin 1994; Bowerman & Choi 2001, 2003; Choi & Bowerman 1991; Özçalışkan & Slobin 1999; Slobin 1996, 2004) show that children are attuned to the structure of their mother tongue from a very early age onward. For example, children learning English or German (satellite-framed) express Manner information more systematically than those learning French (verb-framed) (Hickmann 2010; Hickmann & Hendriks 2010; Ochsensbauer 2010). Such findings suggest that they are not just guided by their conceptual spatial development, but also by the properties of their language.

Given such findings, the question arises whether language-specific factors have implications for cognitive functioning beyond language use. In this respect, some hypotheses revisiting the Whorfian view have put forth the idea that some aspects of linguistic systems may have deeper implications for cognitive organization. Thus, Slobin (1996) has proposed that languages affect attentional processes by filtering incoming information, thereby making some aspects of reality cognitively more salient and more accessible than others. As a result, children learning their native language also learn a particular way of ‘thinking for speaking’.

1.3. Relating gestures and speech in language acquisition

A further vexing question is whether speech alone can reveal everything about emerging representations or whether children also exploit other vehicles of meaning and express different information in gestures. Some studies suggest that gesture and speech develop in parallel (e.g. Bates et al. 1979; Capirci et al. 1996; Nicoladis et al. 1999). For instance, Nicoladis et al. (1999) find that children produce more gestures with increasing age and greater MLU. However, gestures also seem to foreshadow or spearhead language development by expressing alternative information to speech in development. A number of studies find that children at the one-word stage who express complementary information in speech and gesture (e.g. point to ball and say *Daddy*) are more likely to progress to the two-word stage in an imminent phase than are children whose speech-gesture convey the same information (e.g. point to ball and say *Ball*; see Goldin-Meadow & Butcher 2003; Özcalışkan & Goldin-Meadow 2005). Studies also show that children are often able to express information about equivalence and maths in gesture before they are able to talk about them accurately (e.g. Alibali & Goldin-Meadow 1993; Pine et al. 2007).

Interestingly, in the domain of motion, two studies find evidence of gradual development towards adult-like gesturing in Spanish, Mandarin Chinese, English, and Turkish (McNeill 2005; Özyürek et

al. 2008). All children in these studies tend to produce more Manner and Path separately than adults who instead conflate Manner and Path in one gesture. For Turkish and English, Özyürek et al. (2008) find that children only differ cross-linguistically at age nine, whereas younger children all tend to express Manner and Path separately independently of the language spoken. These findings suggest that target-like gesture patterns take longer to establish than target-like speech. They therefore raise questions about the extent to which children's representations of events are adult-like in production. That said, the studies are not equally detailed in their descriptions of concomitant speech or in the granularity with which the speech-gesture relationship is examined. In investigating whether speech and gesture express similar information, studies often compare the meaning expressed in gesture to the one expressed in an entire clause in speech. However, to get a firmer grasp of precisely how children learn to express events in their native language, particularly whether they express similar or different information in speech and gesture and/or whether gestures are recruited to compensate for speech, more fine-grained temporal analyses of exactly co-occurring speech is necessary as a complement to clause-based analyses. It is fair to say that the developmental trajectory towards language-specific speech and gesture as reflections of language-specific means of representing events remains poorly understood.

The present study examines whether children and adults target the same motion information by examining speech and gesture in parallel developmentally and cross-linguistically. Specifically, we ask (1) which meaning components are expressed in speech and gesture, and (2) to which extent speech and gestures are co-expressive both at the clause level (whether the same information is expressed in the clause) and at the level of precise temporal alignment between speech and gesture (whether speech co-occurring exactly with a gesture is co-expressive).

The following predictions can be made regarding language differences: (1) Given the properties of French and English, event descriptions should show a French focus on Path in speech and gesture, but an English focus on Manner and Path in speech as well as possibly in gesture. The following companion developmental predictions can be made: (2) If children operate with adult-like event representations, then they should pattern like adults in speech and gesture within a given language group and same-aged children should therefore differ across languages. (3) If, in contrast, children differ from adults, they should differ from them in speech, in gesture, or in both. Same-aged children may then be similar crosslinguistically, for example when their initial expression is based on language-independent concepts rather than on linguistic input. (4) Children may differ from adults both within a given language and crosslinguistically. That is, French and English children may differ from adult speakers within each language and they may also differ from each other. A final prediction concerns co-expressivity: (5) Children's gestures and speech may be co-expressive even if they are not adult-like. This would suggest a speaker-internal consistency in event representations. Alternatively, if gestures are precursors to speech, children's speech and gesture may express different information,

such that gestures foreshadow or present analogue representations independent of language. In particular, French children may be expected to express Manner in gesture but not in speech.

2. Method

2.1. Participants

The data used in this paper consist of a sub-set of a larger corpus. The combined speech and gesture analyses below focus on three age groups: adults and two groups of children (four and six years) selected among the youngest ages available in the corpus. The analyses are based on speakers who provide more than four gestures during the task. This cut-off point was chosen because the corpus shows a natural divide between speakers who provide four or fewer vs. eight or more gestures, and given that generalizability is very difficult with fewer than four gestures per individual. This selection procedure yielded groups of three speakers for each age in each language: three adults for French (one female) and three for English (one female); three four-year-olds for French ($M=4;0$, two female) and three for English ($M=4;1$, three male); three six-year-olds for French ($M=6;0$, two female) and three for English ($M=6;2$, two female). Children were from the South East of England and from a suburb of Paris. Both samples had a homogeneous middle-class socio-economic background. Participants had to be native monolingual speakers. Children also had to fall within a normal developmental range and show no deficits, as established by their teachers. Adults were mostly young adults enrolled at university. In order to establish the generalizability of these data, we summarize below previous results concerning speech (Hickmann, Bonnet & Taranne 2009; Hickmann & Hendriks 2010) in the larger samples (12 subjects per age), with particular attention to the age groups selected for the present paper.²

2.2. Materials

The materials consisted of 12 short animated cartoons in color (about six seconds each) representing voluntary motion carried out by a variety of agents (animals and humans) in different scenes (see Appendix).³ Six involved motion upwards and then downwards in relation to a vertical ground (e.g.

² The full database for the large sample comprises seven age groups (adults, 3, 4, 5, 6, 8, 10 year-olds) in five languages with 24 participants per age in each language, but analyses are only completed so far on groups of 12 subjects. Additional databases cited in the paper were entirely comparable but involved a slightly different experimental procedure (Hickmann et al., 2009).

³ Participants also described other stimuli including caused motion events (e.g. Hickmann & Hendriks 2010), as well as ‘control’ items in which voluntary motion took place in the absence of any background thereby maximizing the salience of manner and making path minimally relevant (e.g., Hickmann et al., 2009).

tree, telegraph pole...). Six others involved crossing a boundary (e.g. street, river...). In both cases manner varied across items (e.g. walk, run, swim...).

2.3. Procedure

Participants were seen individually in a quiet room. They first saw a training item to get used to the experimental set-up. After each cartoon, the experimenter invited them to tell “what happened” for a “naïve” interlocutor. For adults this interlocutor was a fictitious person who would have to tell the story back on the basis of the taped recording. Children were presented to a large doll which was placed behind the screen as part of a game in which they had to “tell her secrets”. Each participant saw all stimuli. Items were semi-randomly ordered in six different sets to which participants were randomly assigned. The entire session was audio-taped and filmed.

2.4. Coding

The data were coded independently for speech, gesture, and speech-gesture co-expressivity. All data were coded by two different coders. Inter-rater reliability was high in all cases (98% for speech, >95% for gesture identification, >91% for gesture content, > 97% for co-expressivity).

2.4.1. Speech

Each response was coded for all devices expressing Manner, Path, or both Manner+Path (hereafter P, M, MP, respectively). A residual category (hereafter Z) also coded rare unclear cases and especially cases in which neither Manner nor Path was explicitly expressed. Speech (and gestures, see below) were coded at two levels indicating: 1) what information was expressed overall at the utterance level (hereafter *global* analyses); 2) which particular linguistic devices expressed this information (hereafter *locus* analyses), particularly in two loci: main verbs versus all other linguistic means, such as particles, prepositions, adverbials, gerunds (hereafter *Verb* vs. *Other*).⁴ Examples (3) to (8) illustrate this coding (locus analysis in square brackets, global analysis in angled brackets). Note that the Z-category coded the absence of any M or P device at a given locus (e.g. as in (6)) as well as devices that merely expressed general locations rather than Path and/or changes of location (e.g. the locative prepositional phrases in (7) and (8)). Example (8) illustrates a Z-response at the global level resulting from the fact that neither locus expresses M or P.

⁴ ‘Other’ devices occasionally included referring expressions denoting the agent with Manner information (e.g. ‘jogger’).

- (3) *The girl skates* [Verb=Manner] *across* [Other=Path] *the lake*. <global=MP>
- (4) *Elle traverse* [Verb=Path] *en faisant la bicyclette* [Other=Manner]. <global=MP>
 Lit. ‘She crosses by doing the bicycle.’
 ‘She crosses cycling.’
- (5) *Il traverse* [Verb=Path] *la route de gauche à droite* [Other=Path]. <global=P>
 ‘He crosses the road from left to right.’
- (6) *Le garçon nage* [Verb=Manner] [Other=Z]. <global=M>
 ‘The boy swims.’
- (7) *La souris monte* [Verb=Path] *sur le pied de la table* [Other=Z]. <global=P>
 Lit. ‘The mouse ascends on the table leg’
 ‘The mouse goes up the table leg.’
- (8) *Il va sur la glace*. [Verb=Z][Other=Z] <global=Z>
 ‘He goes on the ice.’

A further coding procedure was followed when participants produced more than one relevant utterance for a given item, as illustrated in (9) and (10) (selected utterance in bold).⁵ In these cases, for the purposes of statistical analyses, we identified only one utterance to be included in the counts by selecting systematically: 1) the one that expressed the most information (*richness* criterion, e.g. (9)); or 2) if utterances were equally rich, the one that expressed Path, considered to be the most basic and relevant information (*relevance* criterion, e.g. (10)).

- (9) ***L’ours grimpe à l’arbre***. <global=MP> ... *Il monte*. <global=P>
 Lit. ‘The bear climbs.up at the tree.... He ascends.’
 ‘The bear climbs up the tree... He goes up.’
- (10) ***The boy goes across the river***. <global=P> ... *He is sliding* <global=M>

⁵ Multiple utterances were more frequent among adults than among children and slightly more frequent in French than in English. Proportions for the small sample used for the paper were as follows: English 7% for age four; 11% for age six, 18% for adults; French 7% for age four, 17% for age six, 24% for adults). ‘Richness’ was by far the most frequent criterion used. In the residual cases the ‘relevance’ criterion led to selecting Path-utterances for analysis, although M-utterances were also produced within the same response (but see discussion section below).

2.4.2. Gestures

The gesture coding proceeded in two steps. With sound turned off, gestures were identified and strokes (the meaningful part of the gestural movement where the spatial excursion of the forelimb reaches its apex) and poststroke holds (momentary cessations of an ongoing gesture in gesture space) were coded for whether they expressed Manner, Path, or both (P, M, or MP gestures). We defined Path as a single spatial excursion of a forelimb (following Kendon 2004) and Manner as instances of agitated or repeated movement, or grasping handshapes (following Duncan 2005), and MP gestures as conflating the two. A residual category (Z) labeled rare unclear cases or responses in which neither Manner nor Path was expressed in gestures. This procedure left us with 203 gestures in the experimental conditions, as shown in Table 1.

 Insert Table 1 here

We also coded gestures for their degree of co-expressivity with speech, based on two analyses, examining (1) whether the information in a given gesture occurred anywhere in the co-occurring clause and (2) whether it occurred in the speech with which the gesture was exactly temporally aligned. The clause-based analysis allows for comparisons with previous studies of co-expressivity and the second narrower analysis of exact temporal alignment sheds more specific light on issues of potential crossmodal compensation in development. For the latter analysis, we considered a word to co-occur with the gesture if the stroke or hold covered a vocalized syllable, following procedures outlined in McNeill (1992) and Stam (2006). For example, if the stroke covered only *cl* in *climb*, we did not consider the word to be co-occurring with the gesture, but if the stroke covered *cli*, we did.

This procedure resulted in three possible co-expressivity categories, which we defined as outlined in Table 2 and as illustrated below with English adults (gesture strokes are shown within square brackets, superscripts indicate their expressed content): (a) Match or total overlap, when speech and gesture express the same information, e.g. (11); (b) Speech-Add (S-Add) when speech encodes MP and gesture only M or P, e.g. (12); all other cases are (c) Gesture-Add (G-Add), when speech and gesture express complementary information, that is either one modality expresses M and the other P, or speech expresses only M or P and gesture MP, e.g. (13).

 Insert Table 2 here

(11) **Match** (both under clause and exact analysis):

MP speech (clause) with MP gesture exactly aligned (adult)

He [climbed up]^{MP}.

[Gesture: two grasping handshapes moving upward]

(12) **S-Add** (under clause analysis):

MP speech (clause) with P gesture exactly aligned with P speech (adult)

He just jogs like st[raight acro]^Pss the street.

[Gesture: both hands open palms facing each other moving laterally from the speaker's left to right]

(13) **G-Add** (under exact analysis):

MP speech (clause) with MP gesture exactly aligned with P speech (adult)

She starts skating [across]^{MP} the lake.

[Gesture: both hands open palms facing each other moving repeatedly sagittally out from the speaker's body (as if skating)]

2.4.3. Statistical analyses

The dependent variables were proportions of (1) spoken responses of types P, M, MP, or Z per participant; (2) accompanying gestures of types P, M, MP, or Z per participant; and (3) speech-gesture co-expressivity per participant (Match, S-Add, G-Add). The computed mean proportions were arcsine transformed for statistical analysis (Howell 2002) but non-transformed values are reported in tables, figures and text. Data were analyzed with non-parametric statistical tests, specifically Kruskal-Wallis for comparisons of multiple independent samples and Mann-Whitney for comparisons of two independent samples.⁶

3. Results

3.1. Speech in the large sample

Figure 1 summarizes the response distribution (P, M, MP) at the global utterance level in the large sample (Z was very rare at utterance level and therefore not shown in the figure to simplify presentation). French responses show no predominant pattern, but P-responses are very frequent at all ages. MP mostly occurs among adults but is less frequent than P and rare among children, while M is least frequent and practically non-existent among adults. In comparison, MP-responses are strikingly

⁶ Statistical analyses are only presented below for the small sample which is at the center of the current study, given that only a subset of the large sample is summarized here (12 subjects, also see Note 2) and not yet fully analyzed for statistics. As a result, the figures show either percentages (large sample) or mean percentages (small sample).

more frequent in English than in French at all ages, being predominant among adults and quite frequent among children, despite some M- and P-responses.

 Insert Figure 1 here

Figure 2 displays the locus of expression (main verb vs. other devices) for specific information components that are or are not encoded (P, M, MP Z).

 Insert Figure 2 here

French main verbs typically express Path at all ages and to a much lesser extent Manner (children) or both (adults). Most other French devices express neither Manner nor Path (Z), except among adults who sometimes use them to add more Path information. Although many devices do occur outside of the verb in French, they mostly express general locations (ground information without any Path), rather than M or P, and are therefore coded as Z (e.g. (15) and (16)). French P-responses typically contain a Path verb with no information about Manner (e.g. (14)) and M-responses contain a Manner verb with no information about Path (e.g. (15)). French MP-responses are of two types: with upward motion they involve the verb *grimper* ('to climb.up') that simultaneously lexicalizes MP (e.g. (16)); with crossing events they consist of using a Path verb with a gerund or adverbial phrase expressing Manner (e.g. (17)).

(14) *Il y a une dame en vélo qui traverse* [Path] *un passage à niveau qui est ouvert.* (10 years)
 'There is a woman on a bike who crosses a railroad-crossing that is open.'

(15) *C' est un bébé qui marche* [Manner] *sur la rue* [Z]. (4 years)
 'It's a baby that walks on the street.'

(16) *Un écureuil qui grimpe* [Manner+Path] *dans un chêne rouge* [Z]. (adult)
 'A squirrel who climb.ed.up in a red chestnut tree.'

(17) *On voit un homme qui traverse* [Path] *une route en courant* [Manner]. (adult)
 'One sees a man that is crossing a road by running.'

In sharp contrast, English responses at all ages typically express Manner in the verb and Path in other devices (18). Children also encode Manner in the verb with no information about Path (19), a neutral verb with P elsewhere (20), or occasionally a Path verb (21).

- (18) *The cat just ran [Manner] up [Path] the post.* (10 years)
 (19) *He's skating [Manner].* (4 years)
 (20) *He's going up [Path] the stalk.* (3 years)
 (21) *That's a baby and he's crossing [Path] the road.* (3 years)

In sum, the verbalizations in the large sample show that children in both languages tend to express more motion-relevant information with increasing age, resulting in an increase of MP-responses. Nonetheless, English speakers of all ages typically express both Manner and Path, while French responses are more varied, but tend to focus mostly on Path. As a result, MP-responses are much more frequent in English than in French at all ages.

3.2. Speech in the sub-sample

We now turn to the subsample of adults and children that were selected for the present study. Figure 3 shows the types of spoken responses that were elicited in this subsample and Figure 4 further shows the semantic components that were or were not expressed (P, M, MP, Z) in verbs versus other devices.

 Insert Figure 3 and Figure 4 here

French speakers at all ages predominantly produce P-responses without any reliable differences detectable between age groups (four years M 65% SD 8%; six years M 61% SD 16%, adults M 60% SD 7%; n.s.). Conflated MP-responses increase numerically with age (four years M 6% SD 5%; six years M 24% SD 19%, adults M 36% SD 15%), even though this shift does not reach significance. Finally, M-responses decrease significantly with age (four years M 30% SD 7%; six years M 15% SD 5%, adults M 5% SD 2%; $\chi^2(2,9) = 7.2, p = .03$), such that four-year-olds talk significantly more about M only than six-year-olds ($z = -1.96, p = .05$) and adults ($z = -1.96, p = .05$), and six-year-olds also do so significantly more than adults ($z = -1.96, p = .05$). With respect to information locus, verbs predominantly express Path at all ages (four years 74%, six years 69%, adults 69%), and to a much lesser extent Manner among children (four years 24%, six years 19%, adults 9%) and conflated MP information among adults (22%). At all ages devices outside of the verb are not motion-relevant (four years 98%, six years 80%, adults 57%) with the exception of some Path information among six-year-olds (11%) and especially among adults (35%).

In contrast, English speakers predominantly produce MP-responses at all ages with no reliable differences detectable between age groups (four years M 58% SD 22%; six years M 78% SD 21%, adults M 94% SD 5%; n.s.). English adults never talk about M only, whereas English children do so but gradually less with age at least in numerical terms (four years M 22%, SD 22%; six years M 8% SD 9%; n.s.). English children thus differ from adults in this respect, although they do not differ from

each other. A similar trend can be observed for P-responses where, again, no statistical difference was found between age groups (four years M 20% SD 8%; six years M 14% SD 13%, adults M 6% SD 5%, n.s.). With respect to information locus in English, the following pattern can be observed across all ages: most typically, English verbs express Manner (four years 68%, six years 78%, adults 87%) and other devices express Path (four years 78%, six years 92%, adults 96%). Exceptions concern particularly children's responses which express neither M nor P in the verb (four years 32%, six years 20%) or in other devices (especially four years 19%).

In sum, speech in the subsample shows that children in both languages are adult-like from age four on with respect to the predominant pattern, but approach the adult target gradually with respect to the alternative constructions and to the amount of expressed information. These results are entirely consistent with those discussed above for the large sample. At all ages English responses typically express both Manner and Path, whereas French responses are more varied, but most frequently focus on Path.

3.3. Gestures in the sub-sample

Figure 5 shows the distribution of P, M, and MP information expressed in gestures across the language and age groups. A first result shows that French speakers of all ages predominantly produce gestures about P only. Kruskal-Wallis tests for each verb type with age group as the between-subject factor⁷ reveal that French children are adult-like already at the youngest age. Four-year-olds predominantly produce P only gestures (M 70% SD 21.5%), and do not differ from six-year-olds (M 73% SD 21%) or adults (M 89% SD 9%; n.s.). English adults also predominantly gesture about P only (M 81% SD 16%), and although English children produce numerically more P gestures than adults at age four (100%) and less at age six (M 60% SD 31%), there is no statistically significant difference across age groups. In both language groups, children are therefore adult-like in their main gesture preference already at age four.

 Insert Figure 5 here

In contrast, cross-linguistic and developmental differences are found in other types of gestures. French adults *never* conflate M and P in gesture, whereas French children do. There is also a trend for children to do so less with increasing age, although the comparison does not reach statistical significance (four years M 16% SD 14%, six years M 6% SD 10%; n.s.). French children thus differ from French adults although not from each other. English adults sometimes do conflate M and P in gesture (M 15% SD 8%). In contrast to French children, English children *never* conflate M and P at

⁷ Since not all groups produced all gesture types, we could not perform omnibus analyses on all verb types across all age and language groups.

age four, but they do so more frequently with increasing age (six years M 25% SD 24%; difference to adults not significant). In this case, English four-year-olds differ from older children and adults.

3.4. Speech-gesture co-expressivity

3.4.1. Co-expressivity at the clause level

A first sub-analysis explores co-expressivity between gesture and speech in the accompanying clause. The results are shown in Figure 6. French speakers are predominantly co-expressive (Match) at all ages (four years M 70% SD 10%; six years M 64% SD 15%, adults M 69% SD 19%; n.s.). In cases of discrepancy across the speech and gesture modalities, French speakers tend to add information in gesture relative to speech (G-Add). Thus, in such cases they talk about P in the clause but gesture about M. French children are adult-like in this regard also from the youngest age on (four years M 23% SD 12%; six years M 27% SD 11%; adults M 22% SD 15%; n.s.).

In contrast, the English patterns are quite different. English adults are predominantly *non-co-expressive* and prefer to add information in speech (S-Add; M 61% SD 7%), talking about MP in the clause but gesturing about P. There is a trend towards an effect of age in this regard (χ^2 (2,9) = 5.6, p = .06), and adults do this significantly more often than children at four years (M 31% SD 27%; z = -1.96, p = .05) and six years (M 24% SD 16%; z = -1.96, p = .05), who do not differ. The child groups show preferences that differ across ages four and six, as well as from adults. Four-year-olds mainly add information in gesture (G-Add; M 46% SD 47%) and are numerically more likely to do so than six-year-olds (M 10% SD 1%) and adults (M 7% SD 9%), although the group comparison does not reach statistical significance. Six-year-olds are instead predominantly co-expressive (Match) in speech and gesture (M 66% SD 15%). There is a trend towards a group difference (χ^2 (2,9) = 5.4, p = .07) such that six-year-olds are significantly more likely to Match than four-year-olds (M 23% SD 21%; z = -1.96, p = .05) and adults (M 24% SD 8%; z = -1.96, p = .05), who do not differ.

 Insert Figure 6 here

3.4.2. Co-expressivity in exactly temporally aligned speech

Figure 7 shows the second sub-analysis comparing the information expressed in gesture with the information expressed in exactly temporally aligned speech. French speakers remain predominantly co-expressive (Match) at all ages (four years M 73% SD 3%; six years M 72% SD 15%, adults M 66% SD 18%; n.s.). That is, P gestures are aligned with P speech (Path verbs like *monter* ‘ascend’), M gestures with M speech (Manner verbs like *glisser* ‘slide, skid’), and MP gestures with MP speech

(MP verbs like *grimper* ‘climb up’). In cases of discrepancy, French speakers tend to add information in gesture relative to speech (G-Add) at all ages (four years M 24% SD 1%; six years M 15% SD 15%, adults M 17% SD 15%; n.s.).

Interestingly, English speakers are *also* predominantly co-expressive (Match) at all ages (four years M 54% SD 40%; six years M 78% SD 19%, adults M 61% SD 14%; n.s.). They align P gestures with P speech (Path Satellites like *up*), M gestures with M speech (M verbs like *skid*), and MP gestures with MP speech (Manner verbs and Path satellites like *climb up*). In cases of discrepancy, they tend to add information in speech (S-Add), that is to talk about MP in speech and to gesture about P only. There is a developmental effect in this regard (χ^2 (2,9) = 6.01, p = .05) such that four-year-olds (M 38% SD 38%) add information in speech significantly more often than six-year-olds (M 14% SD 13%, z = -1.96, p = .05) and adults (M 25% SD 6%, z = -1.2, p = .05) who do not differ (n.s.).

Insert Figure 7 here

The following examples illustrate the typical patterns observed in each language (superscripts indicate the content expressed by gesture strokes shown within square brackets). The French adult in (22) and four-year-old in (23) produce P gestures that accompany speech expressing P both when the clause level and exact temporal alignment are considered (Match).

(22) French P speech with P gesture (adult, Match)

Il [monte sur]^P l'arbre.

‘He [ascends on]^P the tree.’

[Gesture: right hand open palm facing left moving upwards]

(23) French P speech with P gesture (4 years, Match)

Il est [monté par la br]^Panche.

‘He [ascended the br]^Panch.’

[Gesture: right hand index finger extended moving upwards]

As for English, the adult in (24) (reproduced from (12) above) and six-year-old in (25) produce P gestures that accompany clauses expressing MP in speech but gestures are exactly aligned with the speech segments expressing P only (S-Add).

(24) English MP speech (clause) with P gesture exactly aligned with P speech (adult, S-Add)

He just jogs like st[raight acro]^Pss the street.

[Gesture: both hands open palms facing each other moving laterally from the speaker's left to right]

- (25) English MP speech (clause) with P gesture exactly aligned with P speech (six years, S-Add)

And he skated a[cross to the other]^P side.

[Gesture: flat hand palm down moving laterally from the speaker's right to left]

4. Discussion

This study compared how French versus English adults and children narrated voluntary motion events, focusing on which meaning components (Path, Manner) were expressed in speech and gesture, as well as on the degree of co-expressivity across these two modalities as a function of language and age.

4.1. Speech and gesture during language acquisition

The results concerning speech in the sub-sample are entirely consistent with those previously reported for the larger sample, confirming that our small subsample is representative of the targeted populations. At all ages verbalizations focus on Path in French, but equally on Manner and Path in English. This first result follows from the typological properties of the two languages. Adults rely on those structures that are most prototypical in their language that constrain which information they do or do not express and where they encode it in the utterance. Children's verbalizations are target-like in that they follow the predominant pattern of the adult system. They also increasingly express more information in both languages (Manner and Path, rather than Manner or Path). Although this numerical change does not reach statistical significance in the sub-sample, it might indicate the possible impact of cognitive factors on how much information children encode within a single proposition.

Turning to gesture, French speakers mainly express Path in gesture at all ages and produce *fewer* conflated MP gestures with increasing age: these gestures are produced only by children but never by adults. English speakers also mainly gesture about Path at all ages but, in contrast to French speakers, they gradually produce *more* conflated MP gestures with age, suggesting a shift toward MP conflation: the youngest children never produce these gestures whereas older children and adults do. By and large, children as young as four overall pattern like adults in both language groups in speech and gesture in the dominant preferences, suggesting overall adult-like expressions of events. However, with respect to the alternative constructions, children differ both from the adults within their language and from each other across languages.

The finding that both French and English speaking children predominantly produce Path gestures are consistent with previous suggestions that children prefer 'separated' gestures, that is gestures that express either Path or Manner rather than conflating both (McNeill 2005; Özyürek et al 2008). The additional observation that English children show adjustments towards more conflation in later

childhood is also consistent with previous studies. Özyürek et al (2008) show that children learning English (but not Turkish) develop an adult-like preference for conflating Manner+Path only around age nine although their speech is adult-like much earlier. They suggest that younger children have processing limitations on how to combine semantic elements in clauses crossmodally. However, the present study shows the inverse pattern for French children, who conflate more at age four than older children or adults. McNeill (2005) has suggested that children's gestures may follow a U-shaped developmental curve such that early instances of conflation may be enactments or imitations of actions rather than reflections of linguistic event construal. Although the French pattern could illustrate this phenomenon, this account would predict a similar pattern in same-aged English children. Given the small sample size in this study, the observation must be treated with caution, but it does raise interesting questions about both processing and enactment accounts which require further experimentation.

Turning to the co-expressivity analyses, French speakers are generally co-expressive in speech and gesture both at the clause level and when exactly aligned speech is considered. Moreover, French children are adult-like in this regard already at age four. In contrast, the two analyses show different outcomes for English speakers. At the clause level, the dominant adult pattern is non-co-expressivity showing MP in speech and P in gesture. Four-year-olds have no clear preference and six-year-olds predominantly Match information in speech and gesture. At the level of exact alignment, English adults and children predominantly become co-expressive by aligning P gestures with P components in speech.

The two co-expressivity analyses highlight several issues. First, it is frequently assumed that gestures replace speech in cases where speakers may have expressive difficulties. In the case of motion expression, the expectation might be that children learning French (V-language) may not typically express Manner in verbs and may therefore gesture M in conjunction with deictic verbal expressions (e.g. *comme ça/like that*). Overall, the data do not support this prediction. Children are overall co-expressive in speech and gesture, suggesting that gestures are not used as a compensatory channel. More compellingly, even cases where gestures accompany deictic expressions that explicitly refer to Manner do not support this idea. Examples (26) and (27) illustrate how both French and English children sometimes use a verb (e.g. *grimper* 'climb up' and *hop*) adding such an expression with a gesture. They clearly have the required lexical Manner expression. Furthermore, the concomitant gesture expresses the same Manner information as the verb. This pattern does not suggest that gestures compensate for missing spoken expressive means, nor that gestures are precursors to spoken expressions of motion, but rather that speech and gesture form an integrated expressive system at these ages.

- (26) *Elle a [grimpe comme ça]^{MP}. (4 years)*
 'She [climb.ed.up like that]^{MP}'.

[Gesture: both hands performing repeated grasping movements while moving upwards]

(27) *He was [hopping like that]^M. (4 years)*

[Gesture: moving whole body up and down on chair]

Second, the discrepancy between co-expressivity in clause-based and exact alignment analyses of English merit comment. In particular, the reduced number of MP conflated gestures may seem surprising, given that the English pattern has been depicted as involving gestures that typically conflate MP. However, the observed predominant pattern combining Path gestures and MP speech at the clause level is an instantiation of what McNeill (2000: 54) calls *Manner modulation*. This phenomenon allows English speakers to downplay Manner and/or to highlight Path in speech by gesturing about Path instead (Brown & Gullberg 2008). Arguably, in V-languages speakers can (and frequently do) downplay Manner by simply omitting it, given the tendency to express Manner peripherally (e.g. *en nageant* ‘by swimming’). The situation is different in S-languages where Manner is expressed in the main verb (e.g. *he swam...*) so that one way to shift focus away from Manner in the clause is to gesture about Path. Crucially, backgrounding Manner and/or foregrounding Path is achieved by aligning Path gestures with spoken Path elements. This is exactly what the English speakers do, as reflected in the predominant co-expressivity observed in the exact alignment analyses. The observation that Manner modulation (or alternatively, Path foregrounding) is achieved by an alignment that does not follow the expected pattern for S-languages has not been highlighted before.

Importantly, however, English speakers *also* produce conflated MP speech and gesture, thus highlighting both meaning components, something French adults do not do. English speakers thus have multiple choices and alternative ways of representing motion. This is an important observation because it underscores the fact that English speakers have different options of construing motion by distributing motion information differently cross-modally relative to French. The choice to background or foreground Manner is presumably a pragmatic one, not necessarily guided by linguistic structures per se, but rather by the communicative situation.

This pragmatic dimension of motion expression may help explain the developmental shift in clause-level co-expressivity observed in English but not in French. It may add to the acquisition challenge for children acquiring English who must learn when one choice is more appropriate than the other. Although children are able to make pragmatic distinctions in various domains at relatively early ages (e.g. Clark 1997; Narasimhan & Gullberg 2006), the choice here may be more challenging. In fact, the developmental data for English suggest that foregrounding Manner may be more difficult than backgrounding it. Recall that under the exact alignment analysis, French and English children are as co-expressive as adults already at age four. Aligning Path gestures with Path speech seems to be a default. In contrast, conflating MP in gesture in English develops gradually and is not even the predominant pattern in adults. Foregrounding Manner and Path in gesture thus seems to be done by

choice. The difficulty for acquisition may reside in taking an alternative perspective on a situation by distributing information across modalities, that is by keeping the spoken form constant and modifying its focus through a shift in information in the gestural channel. In addition, it is not clear exactly when and why English adults alternate between the two construals. It will be important in further studies to expand the contexts of use examined to improve our understanding of when native speakers of languages with multiple choices alternate between one or the other crossmodally.

4.2. Language and cognition across child languages: future directions

Beyond these conclusions concerning the relationship between speech and gesture, our findings suggest some further questions to be addressed concerning the deeper impact of language-specific properties on children's event representations. Since our study is based exclusively on production and on small sub-samples, more research is necessary to unravel the implications of motion expressions across languages.

In addition to analyzing larger samples and more age groups, we are presently pursuing two types of analyses in our corpora to provide complementary information towards a better understanding of event representations. First, previous studies indicate an effect of event types on speech (e.g. Hickmann et al. 2009). In this respect, we saw that French MP-responses to upward motion frequently involved a verb (*grimper* 'to climb up') that lexicalizes Path and Manner (e.g. examples (16) and (26) above). Furthermore, children's M-responses in both languages mostly concerned crossing events (e.g. (15) and (19) above). Previous analyses of the French corpus (Gullberg, Hendriks & Hickmann 2008) also show effects of event types on French gesture at all ages. Thus, when French children produce MP gestures (particularly at age four), they do so more often when describing upwards motion, and when they produce M-gestures, they do so more frequently with crossing events. Comparative analyses presently examine whether event types also affect English gesture.

Second, the low frequencies of MP expression in French as compared to English should be interpreted with care. In this respect, further analyses are necessary to interpret multiple utterance responses (also Note 5 above). Clearly, such responses provide some evidence that Manner information can indeed be accessed by French speakers, even though it may be less central and/or 'fuzzier' than Path in their representations as compared to English speakers. Further analyses of such utterances also show a clear developmental progression in 'response architecture'. In particular, 'loose' multi-utterance responses (e.g. *Il court... il traverse...* 'He is running... he is crossing...') are most frequent among the younger French children, but are gradually replaced by more 'compact' and complex responses involving subordination (e.g. *Il traverse en courant* 'He is crossing by running'). Analyses in progress presently aim to determine how information is distributed across speech and gesture as a function of response architecture across languages and ages.

Finally, in order to generalize conclusions, we are presently examining more varied tasks in a wider range of languages. Other methodologies are also necessary to tap the nature and organization of children's internal event representations *beyond* speech and gesture production. Although gestures shed additional light on the semantic components selected for expression, they are nevertheless tightly bound to the linguistic planning that unfolds online during production. The continuation of this research presently includes other methodologies, for example an eye-tracking paradigm which measures attention allocation to visual stimuli and/or non-verbal tasks (visual recall, categorization). It is only by combining such complementary measures of verbal, co-verbal, and non-verbal performance across languages, that we can begin to examine the language-cognition interface during language acquisition.

5. Conclusion

Children and adults within a given language group (English, French) express the same information when they talk and gesture about motion, but developmental progressions also occur in both modalities and languages. Although caution is necessary given the sample sizes, these findings have theoretical implications for both acquisition and gesture studies. They provide no evidence that gestures act as a 'compensatory' channel in relation to speech during childhood, i.e. that children express alternative information in gestures not provided in speech. Instead, the combined speech and gesture analyses suggest that dominant patterns of event descriptions are language-specific from at least age four but that alternative construals continue to develop in later childhood. Furthermore, the results raise questions concerning how adults and children distribute and organize information across speech and gesture, as well as concerning the nature of gesture-speech integration in development and across languages. More detailed analyses in progress and the use of complementary methodologies should improve our understanding of speakers' underlying and emergent event representations.

Acknowledgements

This research was carried out through the program *Ecole et Sciences Cognitives* and the ANR-DFG project *Langacross* with financial support by the French Ministry of Research and by the Max Planck Institute for Psycholinguistics. Special thanks go to Pierre Taranne and to Sylvain Mottet for their help in programming, as well as to Davide Valentini, Arna Van Doorn and Lisa Mårtensson for their help with gesture coding and interrater reliability. We also thank two anonymous reviewers for helpful comments.

REFERENCES

- Alibali, M. W., & Goldin-Meadow, S. (1993). Gesture-speech mismatch and mechanisms of learning: What the hands reveal about a child's state of mind. *Cognitive Psychology* n° 25(4), 468-523.
- Allen, S., Özyürek, A., Kita, S., Brown, A., Furman, R., Ishizuka, T., et al. (2007). Language-specific and universal influences in children's syntactic packaging of Manner and Path: A comparison of English, Japanese, and Turkish. *Cognition* n° 102(1), 16-48.
- Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). *The emergence of symbols. Cognition and communication in infancy*. New York: Academic Press.
- Berman, R., & Slobin, D. I. (1994). Filtering and packaging in narrative. In R. Berman & D. I. Slobin (Eds.), *Relating events in narrative: A crosslinguistic developmental study*, 515-554. Hillsdale, NJ: Lawrence Erlbaum.
- Bowerman, M. & Choi, S. (2001). Shaping Meanings for Language: Universal and Language-Specific in the Acquisition of Spatial Semantic Categories. In M. Bowerman & S. Levinson (eds.), *Language acquisition and conceptual development*, 475-511. Cambridge: Cambridge University Press.
- Bowerman, M. and Choi, S. (2003). Space under construction: Language specific spatial categorization in first language acquisition. In D. Gentner and S. Goldin-Meadow (eds.), *Language in Mind: Advances in the study of Language and Cognition*, 387-428. Cambridge: MIT Press.
- Brown, A., & Gullberg, M. (2008). Bidirectional crosslinguistic influence in L1-L2 encoding of Manner in speech and gesture: A study of Japanese speakers of English. *Studies in Second Language Acquisition* n° 30(2), 225-251.
- Capirci, O., Iverson, J. M., Pizzuto, E., & Volterra, V. (1996). Gestures and words during the transition to two-word speech. *Journal of Child Language* n° 23(3), 645-675.
- Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition* n° 41(1-3), 83-121.
- Clark, E. V. (1997). Conceptual perspective and lexical choice in acquisition. *Cognition*, 64(1), 1-37.
- Duncan, S. D. (2005). Co-expressivity of speech and gesture: Manner of motion in Spanish, English, and Chinese. In C. Chang, M. J. Houser, Y. Kim, D. Mortensen, M. Park-Doob & M. Toosarvandani (Eds.), *Proceedings of the 27th Annual Meeting of the Berkeley Linguistic Society* (Vol. General Session and parasession on Gesture and Language, 353-370. Berkeley, CA: Berkeley Linguistics Society.
- 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*, 85-107. Mahwah, NJ: Lawrence Erlbaum.
- Gullberg, M., Hendriks, H. & Hickmann, M. (2008). Learning to talk and gesture about motion in French. *First Language* n° 28, 2, 200-236.

- Hespos, S. J. & Spelke, E. S. (2004). Conceptual precursors to language. *Nature* n° 430, 453-456.
- Hespos, S. J. & Spelke, E. S. (2007). Precursors to spatial language: the case of containment. In M. Aurnague, M. Hickmann & L. Vieu (eds.), *The categorization of spatial entities in language and cognition*. Amsterdam: Benjamins.
- Hickmann, M. (2010). Linguistic relativity in first language acquisition: spatial language and cognition. In M. Kail & M. Hickmann (eds), *Language acquisition across linguistic and cognitive systems*, 125-146. Amsterdam: Benjamins.
- Hickmann, M., & Hendriks, H. (2010). Typological constraints on the acquisition of spatial language in French and English. *Cognitive Linguistics* n° 21(2), 189-215.
- Hickmann, M., Bonnet, Ph. & Taranne, P. (2009). Motion in first language acquisition: Manner and Path in French and English child language. *Journal of Child Language* n° 36, 4, 705-742.
- Howell, D. (2002). *Statistical methods for psychology* (5th ed. ed.). Pacific Grove, CA: Duxbury.
- Johnston, J. R. & Slobin, D. I. (1979). The development of locative expressions in English, Italian, Serbo-Croatian and Turkish. *Journal of Child Language* n° 6: 531-547.
- Kendon, A. (2004). *Gesture. Visible action as utterance*. Cambridge: Cambridge University Press.
- Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language* n° 48(1), 16-32.
- Mandler, J.M. (1998). Representation. In W. Damon, D. Kuhn & R.S. Siegler (eds), Vol 2, 255-308. New York: Wiley.
- McNeill, D. (1992). *Hand and mind. What the hands reveal about thought*. Chicago: University of Chicago Press.
- McNeill, D. (2000). Imagery in motion event descriptions: Gestures as part of thinking-for-speaking in three languages. In M. L. Juge & J. L. Moxley (Eds.), *Proceedings of the 23rd Annual Meeting of the Berkeley Linguistics Society, 1997* (Vol. General session and parasession on pragmatics and grammatical structure, 255-267). Berkeley, CA: Berkeley Linguistics Society.
- McNeill, D. (2005). *Gesture and thought*. Chicago: University of Chicago Press.
- Narasimhan, B., & Gullberg, M. (2006). Perspective-shifts in event descriptions in Tamil child language. *Journal of Child Language* n° 33(1), 99-124.
- Nicoladis, E., Mayberry, R. I., & Genesee, F. (1999). Gesture and early bilingual development. *Developmental Psychology* n° 35(2), 514-526.
- Ochsenbauer, A.-K. (2010). *The impact of language-specific factors in first language acquisition: The expression of motion in French and German*. Unpubl. PhD. dissertation, Université de Paris 8/Ludwig-Maximilian-Universität München.
- Ochsenbauer, A.-K., & Hickmann, M. (2010). Children's verbalizations of motion events in German. *Cognitive Linguistics* n° 21(2), 217-238.

- Özçalışkan, Ş., & Goldin-Meadow, S. (2005). Gesture is at the cutting edge of early language development. *Cognition* n° 96(3), B101-B113.
- Özçalışkan, Ş., & Slobin, D. I. (1999). Learning how to search for the frog : expression of manner of motion in English, Spanish, and Turkish. In A. Greenhill, H. Littlefield & C. Tano (Eds.), *Proceedings of the 23rd Annual Boston University Conference on Language Development*, Vol. 2, 541-552. Boston: Boston University Press.
- Özyürek, A., Kita, S., Allen, S., Furman, R., Brown, A., & Ishizuka, T. (2008). Development of cross-linguistic variation in speech and gesture: Motion events in English and Turkish. *Developmental Psychology* n° 44(4), 1040-1050.
- Pine, K. J., Bird, H., & Kirk, E. (2007). The effects of prohibiting gestures on children's lexical retrieval ability. *Developmental Science* n° 10(6), 747-754.
- Slobin, D.I. (1996). From “thought and language” to “thinking for speaking”. In J.J. Gumperz & S.C. Levinson (Eds.), *Rethinking linguistic relativity*, 70-96. Cambridge: Cambridge University Press.
- Slobin, D. I. 2004. The many ways to search for a frog: Linguistic typology and the expression of motion events. In Sven Strömquist & Ludo Verhoeven (eds.), *Relating events in narrative: Vol. 2: Typological and contextual perspectives*, 219-257. Mahwah, NJ: Lawrence Erlbaum.
- Spelke, E.S. (2003). What makes us smart? Core knowledge and natural language. In D. Gentner & S. Goldin-Meadow (eds), *Language in mind: Advances in the study of language and thought*, 277-311. Cambridge, MA: MIT Press.
- Stam, G. (2006). Thinking for Speaking about motion: L1 and L2 speech and gesture. *International Review of Applied Linguistics* n° 44(2), 143-169.
- Talmy, L. (2000). *Towards a cognitive semantics*. Cambridge, MA: The MIT Press.

Résumé

Des recherches récentes montrent que les locuteurs adultes expriment le mouvement conformément aux propriétés des langues ‘à cadrage verbal’ ou ‘à satellites’ (Talmy, 2000), et ceci aussi bien verbalement (Slobin 2004) que dans leurs gestes co-verbaux (Duncan 2005; Kita & Özyurek 2003; McNeill 1992). Si ces résultats indiquent l’existence de différences translinguistiques dans les représentations, les implications de telles différences pour l’acquisition du langage sont encore peu comprises. Cette étude examine comment les enfants (à quatre et six ans) et les adultes francophones et anglophones expriment la Trajectoire et la Manière du mouvement par la parole et par la gestualité lorsqu’ils décrivent des déplacements volontaires présentés sous forme de dessins animés. Les résultats montrent que les adultes anglophones fusionnent verbalement Manière+Trajectoire plus souvent que les francophones, dont les productions verbales se focalisent principalement sur la seule Trajectoire. Les deux groupes produisent principalement des gestes renvoyant à la seule Trajectoire,

mais seuls les adultes anglophones produisent des gestes fusionnant Manière-et-Trajectoire. Dès l'âge de quatre ans, les enfants dans les deux groupes suivent le plus souvent le patron adulte de leur langue, aussi bien dans la parole que dans la gestualité, mais ils font également preuve de progressions développementales. Enfin, la parole et les gestes expriment le plus souvent les mêmes informations dans les deux langues et à tous les âges. Néanmoins, lorsque les deux modalités diffèrent, les adultes anglophones produisent typiquement moins d'informations dans les gestes (Trajectoire) que dans la parole (Manière+Trajectoire, phénomène de 'modulation de la Manière'), alors que les adultes francophones expriment des informations complémentaires dans la parole (Manière) et dans les gestes (Trajectoire). La discussion souligne les implications théoriques de telles analyses bi-modales aussi bien pour l'étude de l'acquisition que pour celle des gestes.

APPENDIX

Stimuli used in the experiment

UP/DOWN-items

1. **Mouse:** A mouse tiptoes to a table leg, climbs up the table leg, takes a piece of cheese, slides back down, and tiptoes away.
2. **Bear:** A bear walks up to a tree, climbs up the tree, puts his paw in a bees' nest, climbs back down, stops under the tree to lick the honey off his paw, and walks away.
3. **Monkey:** A monkey walks up to a tree, climbs up the tree, takes a banana, climbs back down, and walks away.
4. **Cat:** A cat runs up to a telegraph pole, climbs a telephone pole to a birds' nest, drops an egg out of the nest, climbs back down, eats the egg yolk, and runs away.
5. **Caterpillar:** A caterpillar crawls up to a stem, nibbles part of a leaf, crawls back down, and crawls away.
6. **Squirrel:** A squirrel runs up to a tree, climbs up the tree to a hole, goes in and out of the hole, runs down, and runs away.

ACROSS-items

1. **Boy-swim:** A boy runs over to a river bank, swims across the river, and runs away on the other side.
2. **Girl-cycle:** A woman cycles to some railroad tracks, cycles over the tracks, and cycles away.
3. **Boy-slide:** A boy runs over to a frozen river, slides across the river on his feet, and runs away on the other side.
4. **Girl-skate:** A girl walks up to a frozen lake with skates on, skates across the lake, and walks away.
5. **Man-run:** A man runs over to a road, runs across the road, and runs away.
6. **Baby-crawl:** A baby crawls on a pavement, crawls across a cross-roads, and crawls away on the other side.

	4 years	6 years	Adults	Total
French				
N	29	38	36	103
<i>M</i>	10	13	12	11
<i>SD</i>	3	5	1	3
range	8-13	8-17	11-13	8-17
English				
N	29	36	35	100
<i>M</i>	10	12	12	11
<i>SD</i>	1	0	1	1
range	9-11	12-12	11-12	9-12
Total				
N	58	74	71	203
<i>M</i>	10	12	12	11
<i>SD</i>	2	3	1	2
range	8-13	8-17	11-13	8-17

Table 1: Gesture production across age groups and languages:
total number of gestures (*N*), mean number of gestures (*M*), standard deviation (*SD*), and range

Speech	Gesture		
	P	M	MP
P	Match	G-Add	G-Add
M	G-Add	Match	G-Add
MP	S-Add	S-Add	Match

Table 2. Coding of speech-gesture co-expressivity

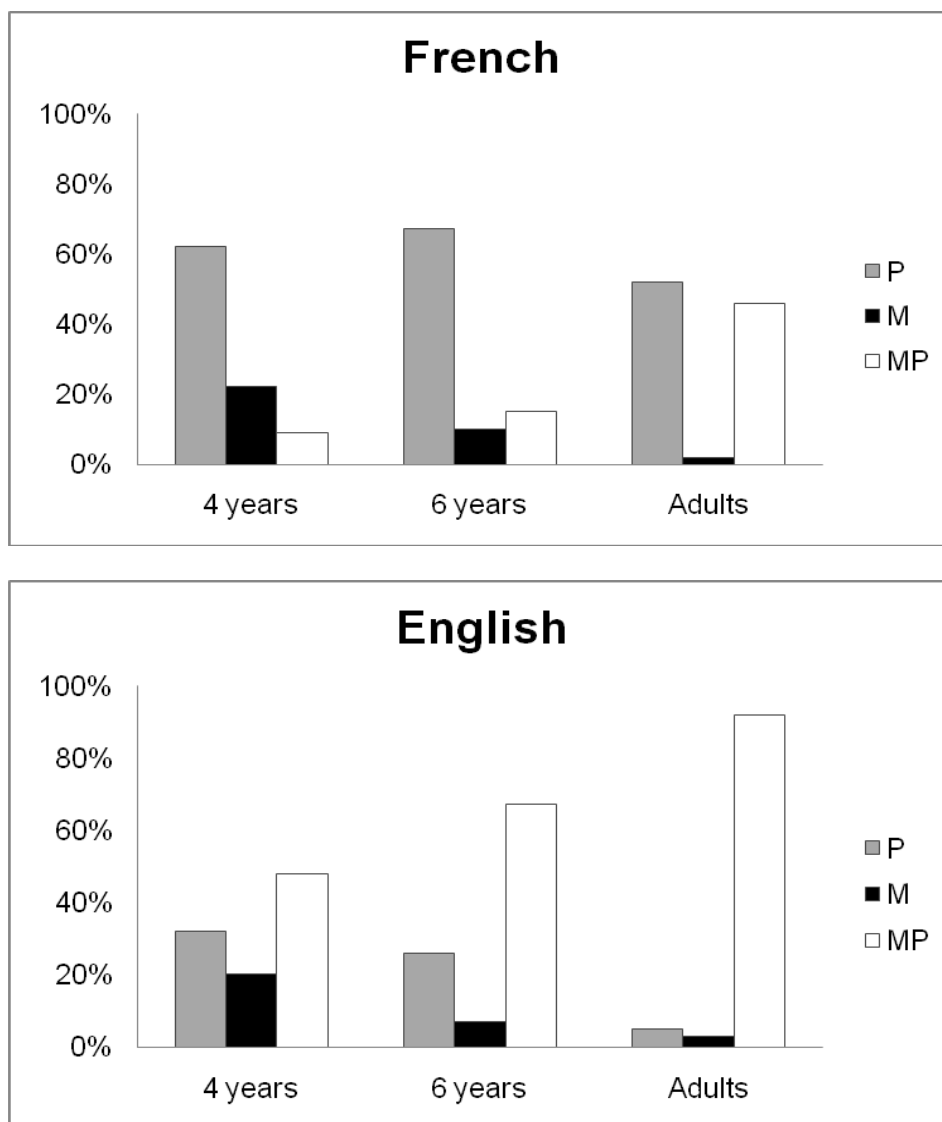


Figure 1. Speech in the large sample: Components expressed at global utterance level
Percentages of responses expressing overall Path only (P), Manner only (M), or both Manner and Path (MP).

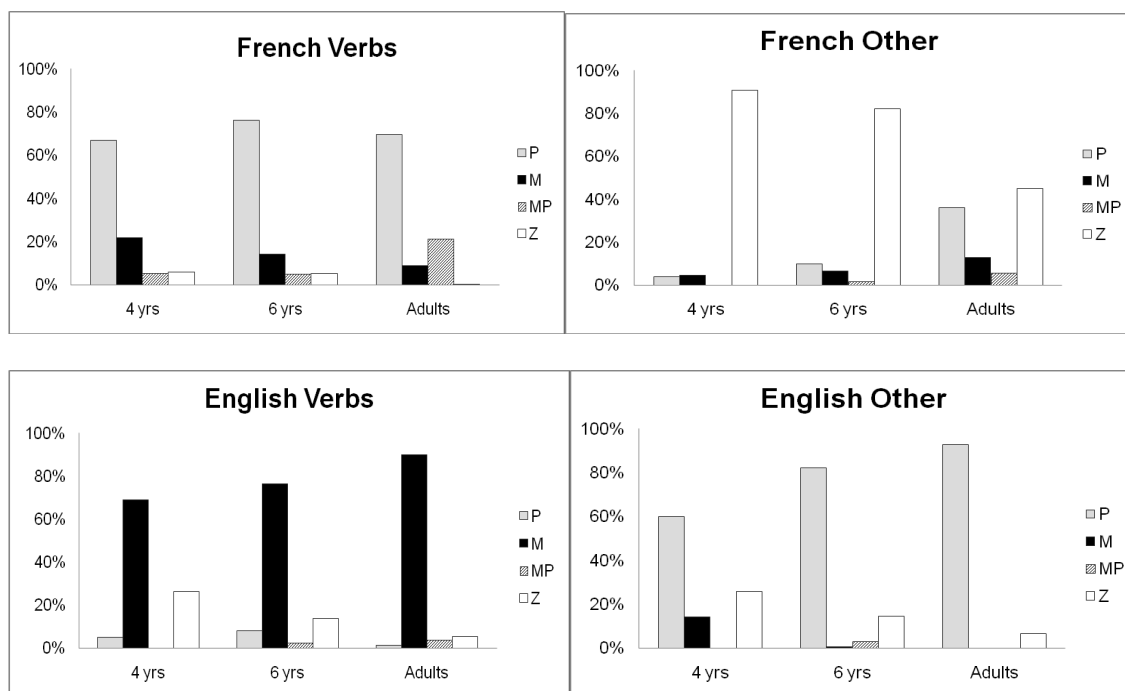


Figure 2. Speech in the large sample: Locus of expressed information

Percentages of verbs vs. other devices expressing Path only (P), Manner only (M), both (MP), or neither (Z).

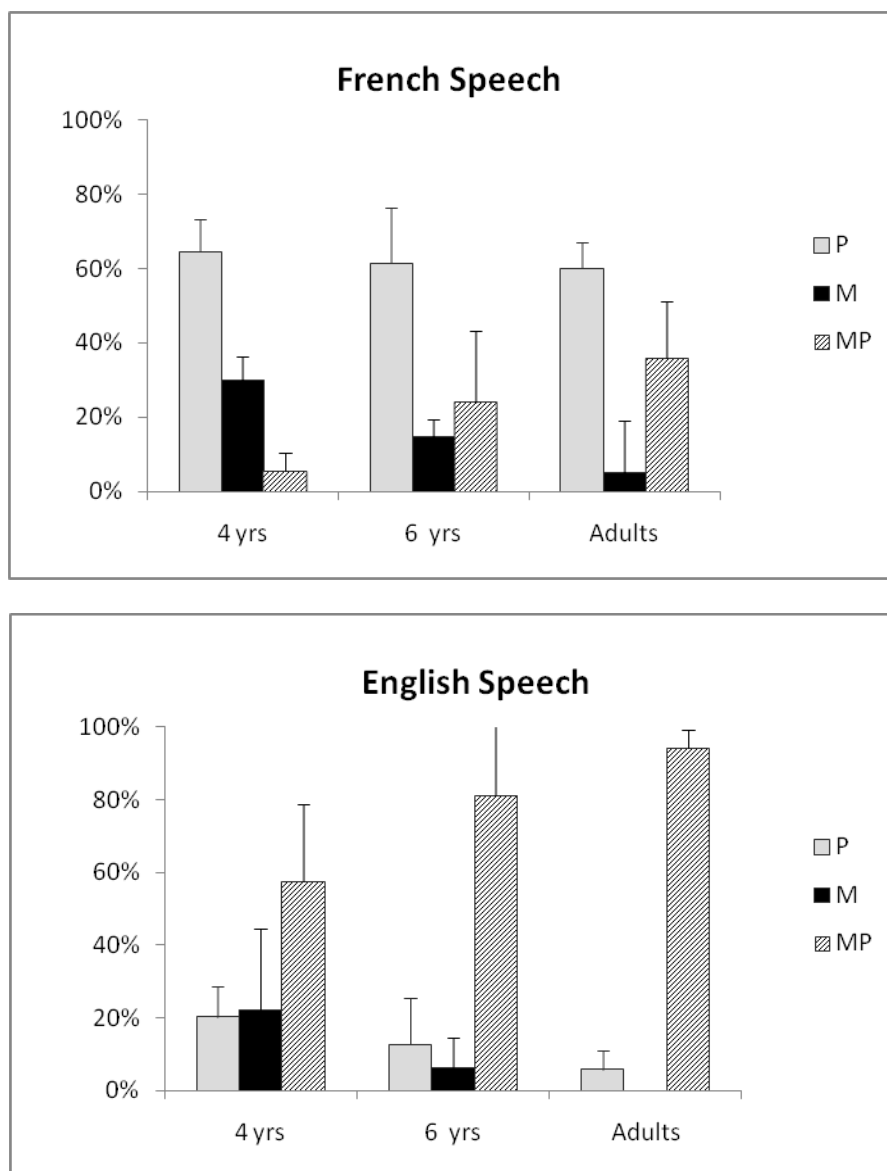


Figure 3. Speech in the sub-sample: Components expressed at global utterance level
Mean proportions of responses expressing overall Path only (P), Manner only (M), or both (MP).

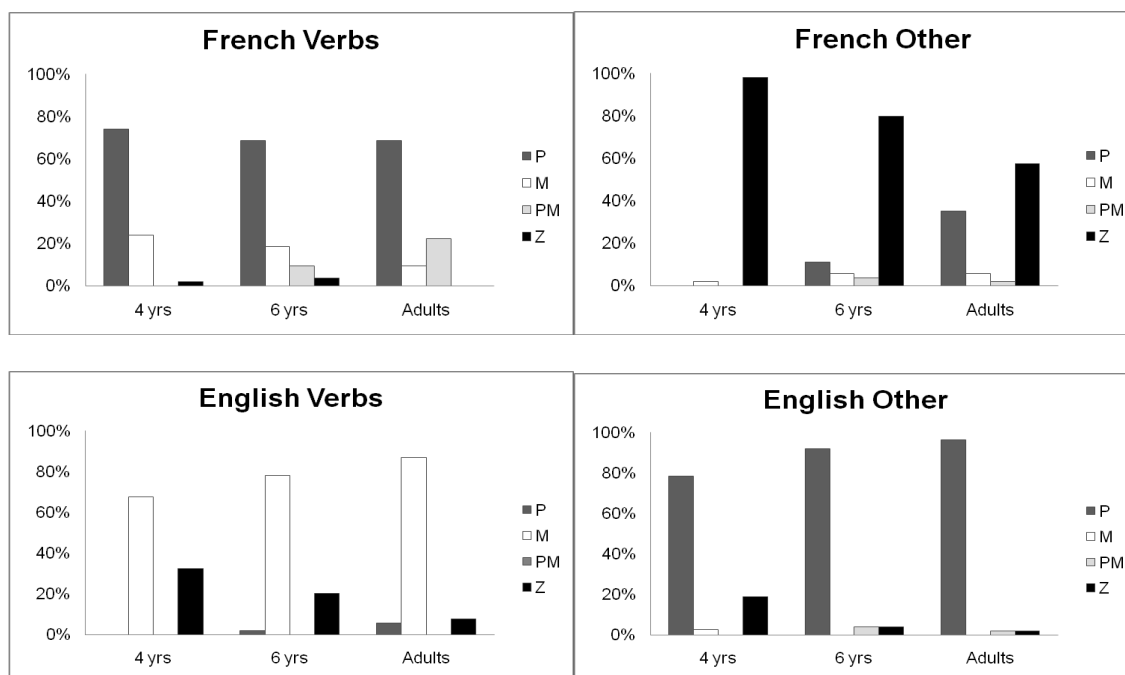


Figure 4. Speech in the sub-sample: Locus of expressed information

Percentages of verbs vs. other devices expressing Path only (P), Manner only (M), both (MP), or neither (Z).

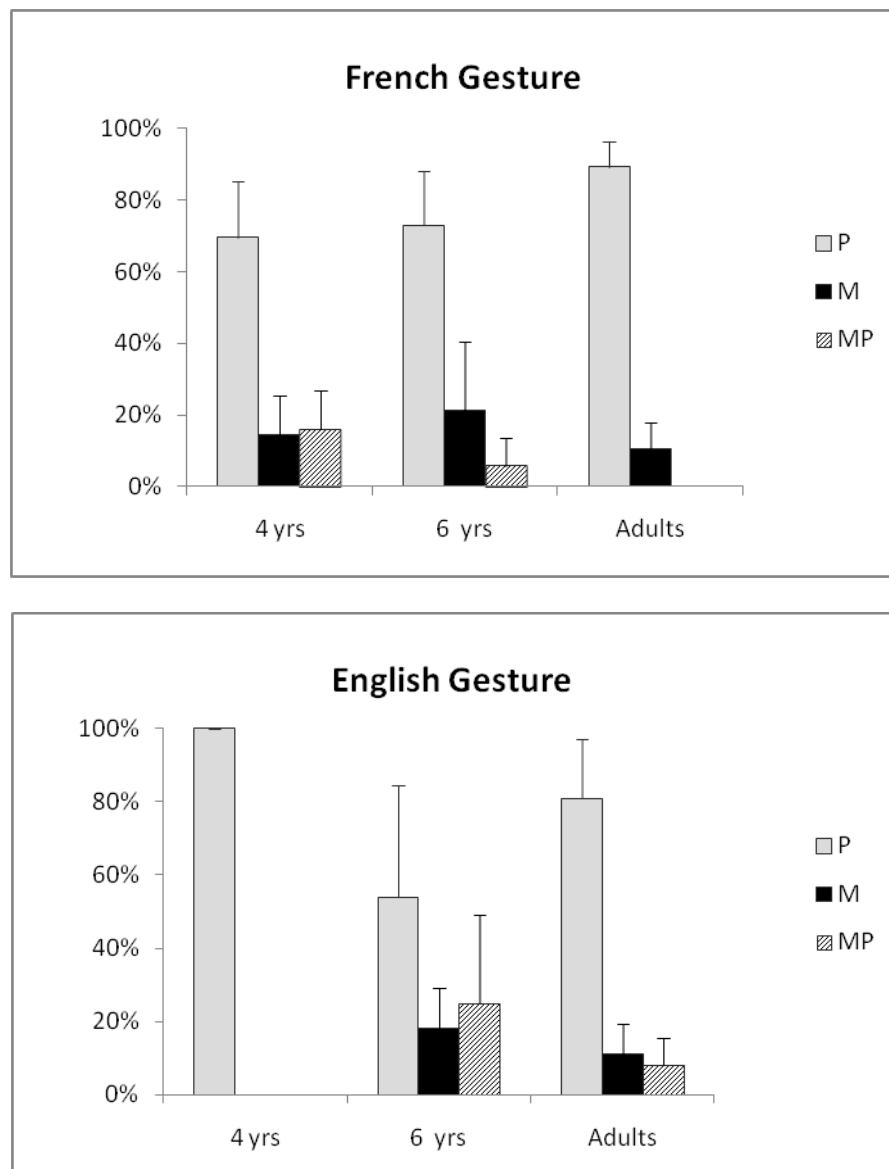


Figure 5. Gestures in the sub-sample: components expressed

Mean proportions of all gestures expressing Path only (P), Manner only (M), or both (MP).

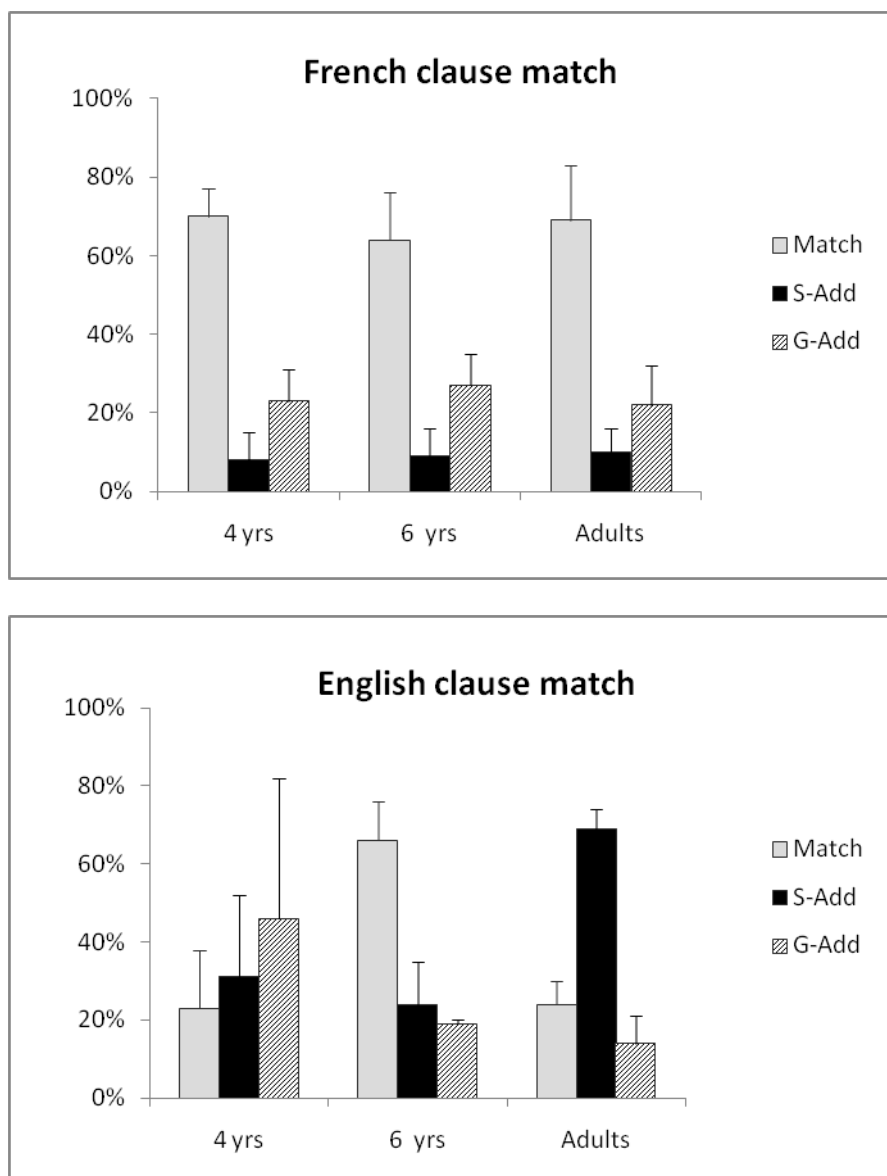


Figure 6. Gesture and speech co-expressivity – clause match analysis

Mean proportions of all gestures expressing Match, Speech-add, or Gesture-add relative to speech

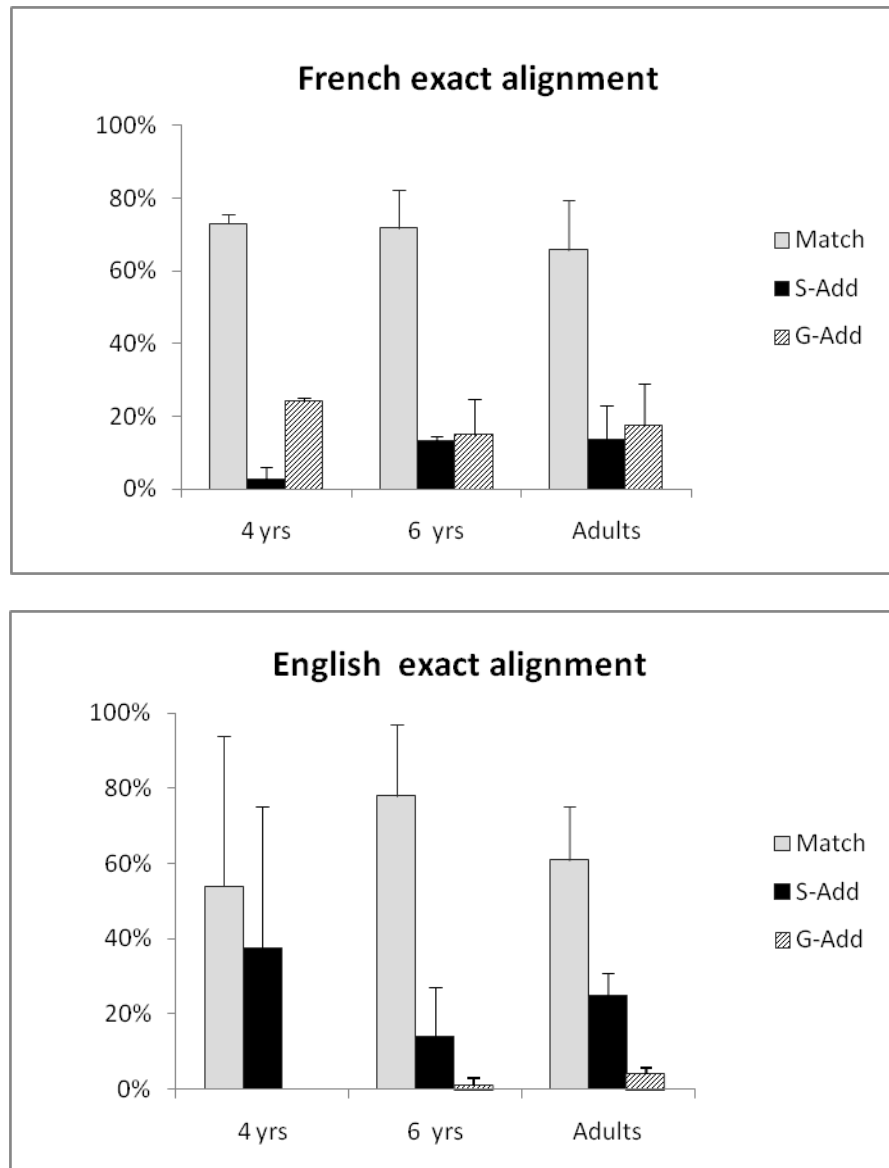


Figure 7. Gesture and speech co-expressivity – match exact alignment analysis
Mean proportions of all gestures expressing Match, Speech-add, or Gesture-add relative to speech