This paper was presented at the 6th Gesture and Speech in Interaction Conference that was held at Paderborn University, Germany from September 11-13, 2019.

To cite this paper:

Fibigerova, K. & Guidetti, M. (2019). Gesture-speech coordination in expression of motion: How far to zoom in to observe semantic synchrony? In: Grimminger, A. (Ed.): *Proceedings of the 6th Gesture and Speech in Interaction – GESPIN 6* (pp. 39-43). Paderborn: Universitaetsbibliothek Paderborn. doi:10.17619/UNIPB/1-806

Gesture-speech coordination in expression of motion: How far to zoom in to observe semantic synchrony?

Katerina Fibigerova and Michèle Guidetti

CNRS-CLLE-University Toulouse 2, France

katerina.fibigerova@univ-tlse2.fr, michele.guidetti@univ-tlse2.fr

Abstract

The present paper contributes to the discussion about coordination between gesture and speech from the semantic and morpho-syntactical perspective. What information is conveyed in co-speech gesture and how that information relates to the content of the co-occurring segment of speech? Does temporal synchronicity imply semantic synchronicity?

We tackled these questions in the context of description of motion events, in terms of combinations of a specific path (e.g. upward, downward, crossing) and a specific manner (e.g. walking, running, flying). We asked whether gesture depicts the same element(s) of motion that speech does and to ensure variability of verbal content we adopted the comparative method involving French and Czech speakers, two languages offering different patterns for expression of motion path and manner.

This paper/talk presents our most recent results that extend our previous studies in this field. After having observed gesture-speech semantic synchrony at the level of 'gesture-proposition' and 'gesture-word', it was time to zoom into individual words and explore the 'gesture-morpheme' level.

1. Introduction and theory

1.1. Gesture-speech synchrony

From the theoretical perspective, we are interested in the type of gesture that is produced during speech (see 'gesticulation' in McNeill 1992). Co-speech gestures are defined as hand and body actions that carry semantic content and co-construct meaning in conjunction with speech (Kendon 2004).

There is evidence the appearance of gestural movements during speech is not fortuitous and that both modalities are semantically synchronized (Kendon 2004; Kita 2000; McNeill 1992). Synchronicity has different aspects. One of them is 'semantic co-expressivity' that describes the fact gesture and speech produced simultaneously share the same reference (i.e. they relate the same thing).

Another aspect is 'semantic redundancy' which indicates the possibility that gesture and speech produced simultaneously also provide the same pieces of information about the shared reference (i.e. they relate to the same thing depicting the same characteristics of it). Since in a multimodal expression, information is distributed and expressed via different modalities/tools/channels which are different and complementary, a bimodal gesture-speech expression benefits jointly from gesture's iconic and holistic qualities and the abstract and analytico-syntactical properties of speech (McNeill 1992, 2005). As a result, some elements of an idea might be expressed in one modality rather than in the other one (see 'information packaging' in Kita 2000).

1.2. Motion in different languages

Motion is understood here as deliberate 'change of placement' (Aurnague 2011) or 'translocation' (Levinson & Wilkins 2006). In the traditional conceptual analysis (Talmy 1985, 2000), motion includes several components: figure that is moving (e.g. a dog), path or direction of motion (e.g. across something), manner or how motion is executed (e.g. by running), and finally ground or the reference point (e.g. the street).

There is evidence that languages vary in lexicalization of information about path and manner (Talmy 1985, 2000), which directly impacts the way speakers of different languages talk about motion (e.g. Slobin 2000, 2004).

In 'verb-framed' languages, path is encoded in the main verb while manner is typically added in gerunds:

1) French:	*descendre*	*l'escalier*	*en courant*
	to.descend	the stairs	by running

In 'satellite-framed' languages, path is encoded in verb satellites while manner is carried in the verb root:

2) English: to run down the stairs

It seems 'verb+gerund' combinations are more complex syntactical constructions—that ask for more robust cognitive treatment—than 'verb+satellite" combinations, which manifests in the fact that speakers of verb-framed languages tend to mention only the path of motion, omitting the manner, while speakers of satellite-framed languages systematically indicate both elements.

1.3. Motion expressed in speech and gesture

When the interest in gesture-speech semantic relation meets the study of cross-linguistic variability in motion expression, the general question often raised up is that of the impact of the latter on the former: When path/manner/both is expressed in speech, is it also present in co-occurring gesture? Are gesture and speech about motion path and manner semantically redundant across languages?

Gullberg, Hendriks and Hickmann (2008) observed gesture-speech relation at 'sentence' level where the considered verbal units co-occurring with speech were either a simple sentence or a complex sentence with subordinate clause. As French native speakers in their study predominantly verbalized and simultaneously gestured about path, it seemed that gesture and speech about motion were semantically redundant (schematically 'path in speech as well as path in gesture'). However, since no comparison to other typologically different language was made, it would be problematic to generalize this conclusion.

A comparative study between English and Turkish speakers was conducted by Özyürek, Kita, Allen, Brown, Furman, and Ishizuka (2008). However, the aim of that study was to explore organization of information inside complex verbal and gestural expressions in which both path and manner were indicated. For this purpose, redundant speech-gesture descriptions of type 'path+manner in speech as well as path+manner in gesture' were elicited in speakers of both languages as a preliminary condition or first step for subsequent observations rather than spontaneously produced by speakers themselves. This is why this particular study does not really fit into our topic.

When Hickmann, Hendriks and Gullberg (2011) extended their previous study (Gullberg, et al. 2008) by comparing French natives to English ones, they concluded that semantic redundancy was impacted by the type of language as well as by the size of verbal unit considered as simultaneous with gesture or, in other words, by the level of observation chosen for analysis of gesture-speech relation. While French speakers produced predominantly redundant gesture and speech (typically 'path in speech as well as path in gesture'), in English speakers, the situation was more complex. At 'gesture – whole surrounding proposition' level, English gesture and speech were semantically mostly non-redundant (typically 'path+manner in speech but path alone in gesture'), which resulted into a statistically significant difference between both language groups. This being said, English gesture and speech turned into mostly redundant at 'gesture – aligned proposition segment' so that the effect of language totally disappeared. The question remained whether the impact of level/unit of observation on semantic relations between gesture and speech was specific to English or it was a characteristic of satellite-framed languages in general.

For this reason, Fibigerova et al. (Fibigerova 2012; Fibigerova, Guidetti, & Sulova 2012; Fibigerova & Guidetti 2018) replicated the study by Hickmann et al. replacing English with Czech. Their French-to-Czech comparison observed first at 'gesture – whole surrounding proposition' level generated results similar to Hickmann et al. The effect of language type manifested through redundancy in French participants (most frequently 'path in speech as well as path in gesture') that contrasted with non-redundancy in Czech group (predominantly 'path+manner in speech but path alone in gesture'). Then, a second analysis was conducted at 'gesture – aligned proposition segment' level. This time, although the proportion of redundant gesture-speech couples increased in comparison to the previously explored level of analysis, the difference between Czech and French speakers was still significant. In conclusion, this last result reported by Fibigerova et al. contrasted

with the result brought by Hickmann et al. showing the absence of difference between English and French speakers at this level of analysis. Both studies considered together seem to reveal deeper intra-typological differences inside satellite-framed languages.

In spite of this conclusion, a doubt remains concerning the stated semantic non-redundancy between Czech gesture and speech about motion and this is why we decided to further explore this point.

2. Present study

2.1. Question and hypothesis

Gullberg, Hendriks and Hickmann (2008) as well as Hickmann, Hendriks and Gullberg (2011) define the segment of proposition considered as simultaneous with gesture in terms of 'word(s)' that are exactly aligned with main gesture stroke. In English, 'gesture-word' level allows to consider the Examples 3 and 4 as cases of redundancy between gesture and speech:

3) A bear climbed up the tree.
+ path gesture aligned with 'up'

4) A bear climbed up the tree.
+ manner gesture aligned with 'climbed'

In Czech language, the situation is more complex. 'Gesture-word' level does help to increase the proportion of redundancy in situations illustrated by Examples 5 and 6 – that are nevertheless less frequent – but not in situations illustrated by Example 7 – that are indeed predominant:

5) *Medvěd* *vyšplhal* *nahoru* *na strom.*
bear up.climbed upwards on tree
+ path gesture aligned with '*nahoru*'

6) *Medvěd* *šplhal* nahoru *na strom.*
bear was.climbing upwards on tree
+ manner gesture aligned with '*šplhal*'

7) *Medvěd* *vyšplhal* *na strom.*
bear up.climbed on tree
+ path gesture aligned with '*vyšplhal*'

At this point, we wonder whether the proportion of redundancy would increase even more if we could deal with situations illustrated by Example 7. For this purpose, we introduce a third level of observation of semantic relation between gesture and speech that we call 'gesture-morpheme' level. Without any aspiration to explore every single gesture-morpheme combination, we use the term of 'morpheme' only as tool that will help us to 'separate' verbal prefix from verbal root so that we could consider them as two different units, each encoding a single element of motion (either path or manner). Thus, we would be able to identify the finest cases of semantic redundancy as shown in Example 8:

8) *Medvěd* *vyšplhal* *na strom.*
bear up.climbed on tree
+ path gesture aligned with '*vy-*'

Our aim is to see whether the difference between Czech and French speakers—after being strong and significant at 'gesture-proposition' level and less strong but still significant at 'gesture-word' level—will finally disappear at 'gesture-morpheme' level. Will we obtain the same effect as Hickmann et al. did in their English-French study after having zoomed in a lower level of observation? We formulate the following hypothesis: the tighter/stricter definition of temporal alignment between gesture and speech, the more semantic synchrony between the two modalities emerges.

2.2. Methodology

Our study is based on 24 French and 24 Czech native monolingual speakers, all young adults (20-35 years old), mainly students, living in their respective countries. The data were collected during individual sessions of watching and narrating short video clips showing 50 different motion events

represented as combinations of specific path and manner. The video clips were several seconds long animated stories with the same structure (a character arrives, realizes the target motion and leaves) that ware created especially for different motion event studies (Allen et al. 2007; Fibigerova 2012; Fibigerova, Guidetti, Šulová 2012; Hickmann 2006). The filmed narrations were then transcribed and annotated using ELAN Linguistic Annotator (10% of our data was annotated by two independent coders).

Firstly, we selected all sentences related to each target motion and all motion related iconic gestures produced during those sentences. Secondly, each gesture was coded according to which element of motion it expressed: a) path alone, b) manner alone, c) both. Thirdly, we identified the segments of speech – one or more morphemes – that were exactly aligned with the main gesture strokes (i.e. the meaningful part of a hand/body movement). Fourthly, each speech segment simultaneous with gesture was coded according to which element of motion it expressed: a) path alone, b) manner alone, c) both. Finally, each gesture-speech couple was annotated as 'semantically redundant' if one of the three situations held: 1) path alone in speech as well as in gesture, 2) manner alone in speech as well as in gesture, 3) both path and manner in speech as well as in gesture. Otherwise, the couple was labeled 'non-redundant' (e.g. both path and manner in speech but path alone in gesture).

2.3. Results

After coding, we proceeded to the comparison of the mean proportions of semantically redundant 'gesture-morpheme' couples produced in both language groups (see Figure 1). The frequency of redundant cases was much higher in French group (M = .73, SD = .121) than in Czech group (M = .26, SD = .222). Since our data were asymmetrically distributed, we used non-parametrical Mann-Whitney U test that confirmed the significance of the found difference (z = -5.217, p <.000).

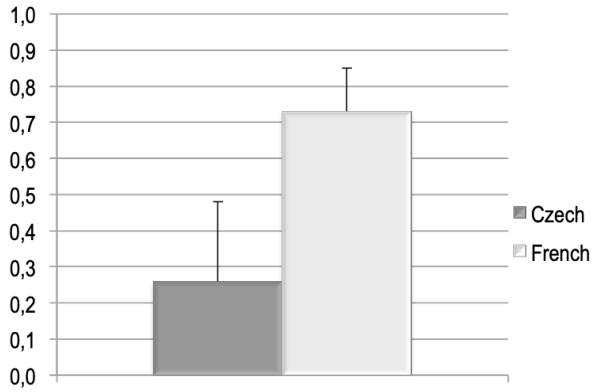

Figure 1. Mean proportions of semantically redundant 'gesture-morpheme' couples.

3. Discussion and conclusion

Contrary to our expectations, we did not obtain similar results to those reported by Hickmann, Hendriks and Gullberg (2011). The difference between Czech and French speakers remains significant even at 'gesture-morpheme' level. Our hypothesis, according to which the tighter/stricter definition of temporal alignment between gesture and speech, the more semantic synchrony between the two modalities emerges, has been only partially confirmed. In Czech (and perhaps in Slavic languages?), semantic redundancy increases slightly at 'gesture-word' level but it decreases again with 'gesture-morpheme' level.

As speakers of satellite-framed languages gesture mostly about path alone, we were the most interested in whether path gestures will be synchronized with path satellites. In spite of very similar situations in English and Czech, the fact that English particles are independent elements placed after the verb while Czech prefixes are bound morphemes placed in front of the verb might be the origin of the different findings reported for these two languages (see Dewell 2011 for analyses of 'prefixed verbs' vs. 'particle verbs' in German).

When 'gesture-speech-mind' unity is concerned, the literature typically mentions two semantic levels: 1) gesture stroke and simultaneous word(s) to express a concept or/and 2) gesture stroke and simultaneous proposition to express an idea (see 'growth point' McNeill 1992, 2005). A bound

morpheme might not be able to function as a sufficiently solid semantic unit so that gesture could be semantically synchronized with one it, independently of the rest of the word. This might also be related to different cognitive processing of information at morpheme vs. word level (see e.g. Giraudo & Voga 2013 for discussion about the place of prefixes inside mental lexicon).

To sum up, we are going back to our very first question. How far to zoom in to observe semantic synchrony in multimodal expression of motion? First of all, we confirm that simultaneously produced gesture and speech are co-expressive, i.e. both of them refer to a given motion. Semantic redundancy between verbal and co-verbal modality—in terms of whether they depict the same conceptual elements of motion (path and/or manner)—however varies with lexico-syntactical specificities of a given language as well as with level of observation/analysis. In French—and probably in other verb-framed languages—redundancy is obvious when we compare a given gesture to the proposition that envelops it. To observe redundancy in English—and maybe in Germanic subcategory of satellite-framed languages—it is necessary to compare a given gesture to the word(s) that is/are produced simultaneously with it. Finally, in Czech—and maybe in Slavic subcategory of satellite-framed languages—even zooming into relation between a given gesture and the semantically meaningful part(s) of words that is/are exactly aligned with it, does not lead to observe any more important semantic redundancy. Gesture and speech produced by Czech speakers are predominantly non-redundant, which makes them different not only from the situation in verb-framed languages in general but also from some other satellite-framed languages.

References

Allen, S., Özyürek, A., Kita, S., Brown, A., Furman, R., Ishizuka, T., & Fujii, M. (2007). Language-specific and universal influences in children's packaging of manner and path: A comparison of English, Japanese, and Turkish. *Cognition, 102*, 16–48.

Aurnague, M. (2011). How motion verbs are spatial: The spatial foundations of intransitive motion verbs in French. *Lingvisticae Investigationes*, *34*(1), 1–34.

Dewell, R. (2011). *The meaning of particle/prefix constructions in German*. Amsterdam & Philadelphia: John Benjamins.

Fibigerova, K. (2012). The impact of language on development of verbal and gestural expression of motion: Comparison between different-aged Czech and French. Unpublished doctoral dissertation, Toulouse: Université Toulouse 2.

Fibigerova, K., Guidetti, M., & Šulová, L. (2012). Verbal and gestural expression of motion in French and Czech. In L. Filipović & K. M. Jaszczolt (Eds.), *Space and time across languages and cultures II.: Language, culture and cognition* (pp. 251–268). Amsterdam & Philadelphia: John Benjamins.

Fibigerova, K., & Guidetti, M. (2018). The impact of language on gesture in descriptions of voluntary motion in French and Czech adults and children. *Language, Interaction and Acquisition*, *9*(1), 99–134.

Giraudo, H., & Voga. M. (2013). Prefix units within the mental lexicon. In N. Hathout, F. Montermini,
& J. Tseng (Eds.), *Morphology in Toulouse: Selected proceedings of Décembrettes 8* (pp. 61-78), Munchen: Lincom Europa.

Gullberg, M., Hendrix, H., & Hickmann, M. (2008). Learning to talk and gesture about motion in French. *First Language*, *28*(2), 200–236.

Hickmann, M. (2006). The relativity of motion in first language acquisition. In M. Hickmann & S. Robert (Eds.), *Space in languages: Linguistic systems and cognitive categories* (pp. 281–308). Amsterdam & Philadephia: John Benjamins.

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.

Kendon, A. (2004). *Gesture*: *Visible action as utterance*. Cambridge: Cambridge University Press.

Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.), *Language and Gesture* (pp. 162–185). Cambridge: Cambridge University Press.

Levinson, S. C., & Wilkins, D. (Eds.) (2006). *Grammars of space: Explorations in cognitive diversity*. Cambridge: Cambridge University Press.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.

McNeill, D. (2005). *Gesture and Thought*. Chicago & London: University of Chicago Press.

Özyürek, A., Kita, S., Allen, S., Brown, A., Furman, R., & Ishizuka, T. (2008). Development of cross-linguistic variation in speech and gesture: Motion events in English and Turkish. *Developmental Psychology*, *44*(4), 1040–1054.

Slobin, D. I. (2000). Verbalized events: A dynamic approach to linguistic relativity and determinism. In S. Niemeier, & R. Dirven (Eds.), *Evidence for linguistic relativity* (pp. 107–138). Amsterdam & Philadelphia: John Benjamins.

Slobin, D. I. (2004). The many ways to search for a frog: Linguistic typology and the expression of motion events. In S. Stromqvist & L. Verhoeven (Eds.), *Relating events in narrative: Typological and contextual perspectives* (pp. 219–257). New York: Lawrence Erlbaum Associates.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical form. In T. Shopen (Ed.), *Language typology and syntactic description* (vol. 3): *Grammatical categories and the lexicon* (pp. 57–149). Cambridge: Cambridge University Press.

Talmy, L. (2000). *Towards a cognitive semantics*. Cambridge, MA: MIT Press.