Coordinating eating and languaging: the choreography of speech, sign, gesture and action in family dinners

Léa Chevrefils¹, Aliyah Morgenstern¹, Pauline Beaupoil-Hourdel², Diane Bedoin³, Stéphanie Caët⁴, Claire Danet⁵, Charlotte Danino¹, Sophie de Pontonx⁶, Christophe Parisse⁶

¹Univ. Sorbonne Nouvelle

²INSPE de Paris

³Univ. Rouen Normandie

⁴Univ. Lille, CNRS, UMR 8163 - STL - Savoirs, Textes, Langage, F-59000 Lille, France

⁵CNRS-LISN

⁶Univ. Paris Nanterre - CNRS

 $leach evrefils {\tt @gmail.com,\ aliyah.morgenstern} {\tt @sorbonne-nouvelle.fr}$

Abstract

In this study, we analyze one French signing and one French speaking family's interaction during dinner. The families composed of two parents and two children aged 3 to 11 were filmed with three cameras to capture all family members' behaviors. The three videos per dinner were synchronized and coded on ELAN. We annotated all participants' *acting*, and *languaging*.

Our quantitative analyses show how family members collaboratively manage multiple streams of activity through the embodied performances of dining and interacting. We uncover different profiles according to participants' modality of expression and status (focusing on the mother and the younger child). The hearing participants' co-activity management illustrates their monitoring of dining and conversing and how they progressively master the affordances of the visual and vocal channels to maintain the simultaneity of the two activities. The deaf mother skillfully manages to alternate smoothly between dining and interacting. The deaf younger child manifests how she is in the process of developing her skills to manage multi-activity. Our qualitative analyses focus on the ecology of visual-gestural and audio-vocal languaging in the context of co-activity according to language and participant. We open new perspectives on the management of gaze and body parts in multimodal languaging.

Index Terms: Family dinners; French, French sign language (LSF); co-activity; *languaging*

1. Introduction

Family dinners grounded in commensality are a collective ritual that plays a key role in family members' identity and constitutes an inherent part of their cultural heritage. Those shared moments of everyday life present a perfect opportunity to study how situated multimodal language and interactive practices are transmitted to and used by children in order for them to construct meaning. Because the subtle interweaving of these practices while eating fully engages the body, our family dinner project highlights the semiotic differences between parents and children using a spoken language, *French*, and a sign language, *Langue des Signes Française* (LSF). Ethnographic methods are used to collect dinnertime data. Qualitative analyses are combined with quantitative methods based on manual annotations. Our aim is to capture the multimodal forms and

functions of situated practices, as they shape and propel the dynamics of family life.

In family dinners, language practices can be analyzed as they occur in real life and real time in the framework of multiparty interactions and multiactivity (Haddington et al. 2014) to capture the multiple deployments of the embodied behaviors of speakers and signers. In this study, we analyze all the members of a French signing and a French speaking family's management of actions, gaze, speech, sign and gestures. We focus on children's socialization to the finely-tuned coordination and in situ organization of the joint activities of conversing and dining that fully engage the same body components. Our aim is to show how family members collaboratively manage the accomplishments of multiple streams of activity and coordinate their temporal organizations through the embodied performances of dining and interacting (Goodwin, 1984).

We first present our theoretical framework, our data and methods, our quantitative results on the coding of one LSF signing and one French speaking family and then qualitatively illustrate the similarities and differences in the orchestration of body parts while eating and *languaging*.

2. Theoretical framework

Our theoretical framework combines language socialization, cognitive grammar, gesture studies, anthropology, interactive and multimodal approaches to situated languaging across languages, modalities and cultures. We borrow the term languaging to refer to multimodal language use - "linguistic actions and activities in actual communication and thinking" (Linell, 2009: 274) expanding the term to include speaking, signing and gesturing. We study how children's socialization to a variety of modes of expression in their daily experiencing (Ochs, 2012) through dinners shapes language socialization. The framework of Cognitive Grammar provides a means of taking into account all semiotic resources as a consequence of the usage-based (Langacker, 1988) nature of the theory. The theory allows for linguistic signs (in the Saussurian sense, 1916) to be multimodal to varying degrees, based on the extent of schematization and entrenchment. Adults have the skill to coordinate the semiotic resources at their disposal, varying the use of "the scope of relevant behaviors" as needed (Cienki, 2012), adjusting to the context of interaction, the activity, the age and identity of the interlocutor, the time of day, etc. Those skills are transmitted to children through languaging in a

variety of situations. Each language provides a certain set of options for the grammatical encoding of characteristics of objects and events. If children are "guided in how they choose to talk about experience by the most available grammatical means provided by their native language" (Slobin, 1987: 443) as they are "thinking for speaking", the same could apply to "thinking for signing" or "thinking for speaking in several languages". Languaging might thus not be solely relative to languages and cultures, but also to the mode of expression as we embody mental construals. We use the framework of construction grammar (Goldberg, 2006; Tomasello, 2003) and apply it to "multimodal constructions" (Kendon, 1988; Andrén, 2010; Morgenstern, 2014) and are in the process of enlarging the framework to include certain conventionalized movements and manipulations of objects, in line with Mondada's research methods (2016, 2019). We first applied this new approach to language and its use to longitudinal and mostly dyadic data in the past years (Morgenstern & Parisse, 2017; Morgenstern et al. 2021; Morgenstern, 2021; Morgenstern & Goldin-Meadow, 2021; Beaupoil-Hourdel & Morgenstern, 2021). For several authors, gestures and speech are part of the same system (e.g., McNeill, 1992). In parallel with Darwin's observations (1872), Bolinger (1983) highlights that gestures are coupled with intonation and display the same ascending and descending movements. Balog and Brentari (2008) showed that children coordinate their verbal and non-verbal behaviors at the temporal and directional levels as early as the first word period. Following Boutet (2018), we have been analyzing the bodies of our participants as both the support (the instrument) and the substrate (which constitutes and structures) of what we call interlanguaging. We are now applying this framework to family dinner situations (Morgenstern et al. 2021).

Detailed analyses of dinner conversations, inspired by language socialization theory, have become a rich source of data for the study of social practices and conversation (Blum-Kulka, 1997; Ochs, Pontecorvo & Fasulo, 1996; Mondada, 2009). Eating and conversing are activities constitutive of the family (Ochs & Kremer-Sadlik, 2013). Central to this research project is a focus on the language socialization practices of families (Ochs, 1988; Ochs & Schieffelin, 1984; Schieffelin & Ochs, 1986).

In this study, we observe and analyze participants who are engaged in dining together and focus on how the language they use, LSF or French, the semiotic resources at play and the body parts involved in their co-activity, shape their *interlanguaging*.

3. Data and Method

3.1. Data

Our previous research prompted us to build a multimodal, multi-dimensional corpus allowing for a multi-track annotation scheme. Our work on sign languages (Morgenstern, 2014; Blondel et al., 2017 among others) has reinforced our idea of analyzing actions, gestures and the role of the body in order to understand the functioning of both sign languages and so-called "spoken" languages. This is also why we decided to create a corpus composed of two linguistic and cultural contexts within the same country, in order to better understand them, with their differences and their common features. Our aim is to analyze the productions and behaviors of families a) for which LSF (Langue des Signes Française) is the main language of communication and b) for which French is the main language of communication.

We collected data from families composed of two adults and two to three children aged 4 to 10 years old. For this first study, in order to test our time-consuming coding scheme we restricted our analyses to one family in which the four members are deaf and use LSF (which we call the *LSF-family*) and one family in which the four members are hearing and use French (which we call the *FRA-family*).

Our corpus is based on situations of spontaneous use of visible and audible *languaging* (consisting in speech, sign and gesture), in the context of an everyday activity. We chose to record French family meals because they are situations of great social and cultural importance; interactions are multiple and not limited to two participants; parents and children are associated in shared situations; these moments of interaction are "ordinary" enough not to be felt as too intimate to be shared with our team and with the participants' informed consent, in the framework of open-access archiving, with the scientific community in the framework of our research program. Our procedure was approved by the Ethics committee of Sorbonne Nouvelle University.

3.2. Video-recording procedure

After several different trials, we devised a filmic apparatus that enabled us to capture the use of all the semiotic resources at the participants' disposal in our video and audio-recordings. In order to alter the usual flow of the dinner as little as possible, we used the following setup (Fig. 1):

- two conventional cameras equipped with high-quality external microphones and arranged to allow a view from the left of the dining scene and a view from the right of the scene;
- a 360° camera placed in the center of the table and offering a front view of all participants;
 - a 360° sound recorder placed next to the 360° camera.



Figure 1: Filming equipment

Video and audio editing is then done to obtain a temporal alignment of all the media. Extractions from the 360° camera are made to best visualize the faces, gaze and the rest of the body of the participants.

3.3. Coding method

Our research question naturally guides our coding and analyses, in particular:

- -Are there differences according to the language used (French and LSF -including the co-verbal or co-sign gestures) in the amount of co-activity (dining activities and languaging activities?
- Does the amount of co-activity affect the amount of multimodal *languaging* used in each family?

In order to tackle these issues, our coding integrates all gestural-visual and audio-vocal *languaging* (including gestures), all dining and non-dining *acting*. An additional constraint is that most of the analyzed elements are coded independently at the temporal level. Thus, the temporality of gestures, of spoken or signed utterance, of *acting*, even if they can be considered as being at least sometimes interdependent (at the intra-individual and inter-individual level), are coded separately.

The videos were thus synchronized and coded using one of the most common tools used in multimodal analyses, ELAN (https://archive.mpi.nl/tla/elan). We created an annotation scheme, or template in this software, and a coding manual that allows us to code all the parameters indicated above. This template and the coding manual will be provided with the corpus and examples of coding when we give access to the platform completed data on the Ortolang (https://www.ortolang.fr/fr/accueil/). For this study, we annotated all participants' acting, and languaging throughout the two dinners.



Figure 2: CODING scheme in ELAN

Table 1: Presentation of the tiers used in our ELAN template

| Tiers | Explanation |
|----------|--|
| lang-aud | Segmentation of the auditive languaging |
| interloc | Identification of the interlocutor(s) |
| script | Transcription of audible language productions |
| lang-vis | Identification of symbolic gestures and/or sign language production |
| act-vis | Identification of all acting produced and categorization into dining/non dining activity |
| interact | Identification of the person the action is oriented towards |
| gaze | Participant or object gazed at |

4. Quantitative data

One of the major difficulties encountered was how to optimize the results of our coding to provide descriptive or inferential statistics. Indeed, the lack of temporal correspondence between the different elements coded in the ELAN tiers does not allow a multivariate type of structure for analysis and statistics. For example, in order to know if a gaze is associated to an action or to languaging, the elements of the "gaze", "acting" and "languaging" tiers must first be linked before the various features can be counted or measured. We therefore need to relate elements spread over several tiers on the basis of their temporal relationship, and without any dependency relationship. This linking process can be done in two ways which we have both tested. The first method is to use the structured queries implemented in the ELAN software. This method is extremely powerful as many variations in the constraints can be applied to the different tiers. For example, constraints can be applied to temporality (before, after, during, etc.), but also to elements included in the forms that are searched, to the selection of tiers to be explored, etc. The results can be exported in the CSV format into a spreadsheet or statistical software. However, as computation time becomes too long, it is not possible to create constraints on more than two tiers at the same time (we must therefore express a chain of multiple constraints pertaining to two tiers each to manage more than two tiers at the same time). Another method that is more cumbersome to apply but much more powerful and quicker to use is to build a python program that uses libraries allowing the analysis of ELAN files. This method is more powerful because the potential of a programming language is infinite. On the other hand, it is necessary to create a specific program for each query, or to create more complex programs with multiple parameters. In the DINLANG project, we have chosen to use specific programs. They have the advantage over the tool integrated in ELAN of doing the operations quickly in a single command, but the disadvantage of their specificity is that they allow only one type of calculation. For more flexibility of use, they are available as a web service easily accessible to the project members (see https://ct3.ortolang.fr/toolselan/).

5. Quantitative analyses

5.1. Comparison between the family profiles

We will first focus on the time globally spent by families *acting*, *languaging* (which includes sign, speech and gestures) or doing both at the same time. All our results are measured in time and not in number of occurrences (we measured the duration of our annotations in ELAN). The FRA-family dinner lasts 25 minutes and the LSF-family dinner lasts 34 minutes.

For the FRA-family (Fig. 3, left graph) *acting* alone is the predominant category, and amounts to 66% of the duration of the dinner. 24% of the time is used *languaging* and *acting* at the same time, i.e. in co-activity. Finally, 11% is devoted to *languaging* alone. For the LSF-family (Fig. 3, right graph), *acting* alone is also largely predominant since it represents nearly 64% of the dinner, *languaging* comes second with 20% of the dinner duration, and co-activity follows with 16%.

Both families therefore spend a large part of their dinner producing bodily actions that may be meal-related, such as drinking from a glass, or non-meal-related, such as playing with their own cutlery (without using them to eat). The proportion is similar between the two families although the LSF-family spends slightly more time *acting* alone.

However, these results indicate a difference in the way these two families coordinate their language use and their dining or non-dining activities. Indeed, within the FRA-family, we find a lower proportion of *languaging* alone (11%) than of co-activity (24%), which suggests that participants tend to

interact with each other without having to interrupt their *acting* and vice versa. For the LSF-family the rate is the opposite since the time spent on *languaging* alone (20%) is higher than in coactivity (16%), which suggests that the members of this family tend to alternate *languaging* and *acting*. Interestingly enough, in both families, *languaging* (as a mono-activity or a coactivity) represents 35% and 36% of the duration of the dinner, slightly more in the LSF-family than in the FRA-family.

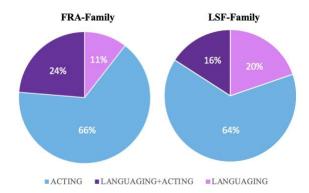


Figure 3: Percentage of acting and of languaging (duration) in each family

To better capture the complexity of these general results, let us focus on the *languaging* and *acting* of two members of each of these two families, the mother and the youngest child, as their profiles differ the most.

5.2. Focus on mothers and younger children

In our family dinners, the mothers often control the flow of the meal and its co-activities by, for example, ensuring that children have eaten enough or by mediating conversations. Conversely, younger children are the most dependent on this supervision and are more likely to engage in practices that require regulation. Young children are also developing their multitasking skills.

We now focus on the *languaging* and *acting* profiles of these two differing participants (Fig. 4).

These data allow us to conduct a targeted inter-family comparison and corroborate our general results: both members of the FRA-family present a higher propensity to co-activity (average at 27,5%) than the LSF-family (average at 12%). The semiotic channel used to communicate thus seems to have a strong impact on the potential for overlapping activities (*languaging* and *acting*) deployed during dinner. Both mothers use language more than the average in their family profile, about one third of the time. The younger children use language less, 25% to 29% of the time. There is slightly less *languaging* in the LSF-child than the FRA-child's profile.

This focus on two family members also allows us to underline that the young children present in both cases a lower proportion of co-activity than their mothers. This shows that the simultaneous orchestration of language and dining practices requires a particular mastery that is probably acquired and developed over the years. Moreover, the gap is less important for the FRA-family than for the LSF-family, suggesting that it is more complex for the youngest child signer to acquire the skill to use sign-language and actions simultaneously as they mobilize the same body segments. However that does not seem

to impact their overall amount of languaging during dinner time.



Figure 4: Percentage of acting and languaging (duration) of mothers and younger children in the two families

We now present a set of qualitative examples to illustrate and refine the analyses presented so far.

6. Qualitative analyses

As we have shown, ChildB in the LSF-family exhibits strong mono-activity: she either produces language; or she is visually engaged in the other participants'responses to her own *languaging* without associated acting; or she is eating. She also shows less attention to her environment than the other members of her family. These behaviors are illustrated in Fig. 5 (video available here): ChildB looks at her plate while eating, and in doing so she cuts herself off from the surrounding discussions. She does not notice the father's wishing them to enjoy their meal whereas her mother and her sister respond to him. As a result, the mother calls her several times by tapping her shoulder, and finally places the source of her request, grated cheese, in front of her eyes, in order to catch her attention.

In the FRA-family, ChildB combines *acting* and *languaging*. However when the mouth is used to do both at the same time, the situation can lead to misunderstandings. In Fig. 6 (video available <u>here</u>), ChildB wants to intervene in the discussion while he has a piece of bread in his mouth. This coactivity is not appreciated by his parents. It is pragmatically incorrect, his speech is not understood and he is immediately reprimanded by his mother.

As far as the mothers are concerned, Fig. 7 (video available here) highlights the fact that the LSF-Mother alternates between sequences of dining activity, and sequences of signed

languaging. This alternation is made necessary by the use of the same body segments, the upper limbs, for both activities.



Figure 5: Mother presents grated cheese to ChildB (Click here to watch the video)



Figure 6: ChildB speaks with mouth full (Click <u>here</u> to watch the video)



Figure 7: LSF signing Mother alternating activities (Click here to watch the video)

Fig. 8 (video available here), shows FRA-Mother talking while helping herself to food; then talking with food in her mouth. Although the mother eats with her mouth full, as we observed for FRA-ChildB, she does so with expertise: she politely hides her mouth and manages to articulate enough to be understood by the other participants. Her different dining activities do not prevent her from conducting co-activity.



Figure 8: French-speaking Mother's fluid co-activity (Click here to watch the video)

7. Concluding discussion

The multimodal nature of language lead us to study how various semiotic systems such as speech, gesture, facial expressions, but also object manipulations, are simultaneously deployed, transmitted and used in family dinners.

This study has enabled us to highlight the fact that the management of multi-activity, utensils, food, actions, gestures and words/signs is different according to the language practiced, but also according to motor and cognitive development. In both our families, expertise in co-activity seems to develop with age and experience. The members of our FRA-family have more access to co-activity as they combine speech and acting. But they also need to be socialized to acceptable co-activity as ostensive chewing and speaking (in our FRA-family) is not culturally acceptable. Co-activity is less accessible to the members of our LSF-family as they predominantly use their upper limbs both for languaging and acting. The younger child is focused on one activity at a time: eating and producing or receiving languaging seems to be mutually exclusive most of the time. But as they develop their competence, signers seem to skillfully and smoothly alternate activities. They manage co-activity as much as possible when necessary thanks to the rich affordances of their two hands and two arms. We also found that overall, the modality used for languaging does not seem to alter the average amount of languaging in the two family dinners but that the lesser experience in the orchestration of activities does affect the amount of languaging produced.

Our results on the ecology of visual-gestural and audio-vocal *languaging* in the context of co-activity according to language and participant are to be considered in the framework of the very modest scale of this study restricted to two families. In order to extend our findings and to capture the specificity of LSF signers and French speakers, we need to collect and analyze more dinners in more families and also focus on two important aspects:

- The use of specific body segments for both *acting* and *languaging* in signers and speakers: the mouth is used for eating/chewing and speaking and the upper limbs (arms and hands) are used for signing, gesturing and dining activities.
- The use of gaze for *interlanguaging* as interlocutors very often gaze at the speaker in vocal languages of course but they have the ability to listen without gazing, whereas gaze on the signer (or having the signer in one's field of vision at least) is necessary for perception and comprehension of sign language.

Our next studies will focus on those specific aspects in our two family dinners then be expanded to all the data collected in the DinLang project. Our aim is to analyze the subtlety of human beings' skillful coordination of the affordances of their bodies to conduct *languaging* and *acting* in multiparty interactions and how children are progressively socialized to orchestrate all the semiotic resources at their disposal according to their cultural and linguistic environment.

8. Acknowledgements

The authors would like to thank the Agence Nationale de la Recherche for its funding of the project.

9. References

- Andrén, M. 2010. Children's Gestures from 18 to 30 months, PHD dissertation, Centre for Languages and Literature, Centre for Cognitive Semiotics, Lund University.
- Balog, H. and Brentari, D. 2008. The relationship between early gestures and intonation. First Language, 28, pp. 141-163.
- Beaupoil- Houdel, P., Morgenstern, A. 2021. French and British children's shrugs. A cross-linguistic developmental case-study of a recurrent gesture. *Gesture* 20:2: 180–218.
- Blondel et al., 2017
- Goldberg, Adele, E. (2006). *Constructions at Work*. Oxford: Oxford University Press.
- Blum-Kulka, S. and Snow, C. 2002. Talking to adults: he Contribution of Multiparty Discourse to language Acquisition. Mahwah, N.J., US: Lawrence Assoicatiates Publishers.
- Bolinger, D. 1983. Intonation and gesture. American Speech. Vol. 58, No. 2: 156-174.
- Boutet, D. 2018. Pour une approche kinésiologique de la gestualité. (Habilitation à diriger des recherches). Université de Rouen-Normandie.
- Cienki, A. 2012. Usage events of spoken language and the symbolic units we (may) abstract from them. In *Cognitive processes in language*, edited by Janusz Badio and Krzysztof Kosecki, 149-158. Bern: Peter Lang.
- ELAN (Version 6.2) [Computer software]. (2021). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Retrieved from https://archive.mpi.nl/tla/elan
- Goldberg, Adele, E. 2006. Constructions at Work. Oxford: Oxford University Press.
- Goodwin, C. 1984. Notes on story structure and the organization of participation, In *Structures of Social Action: Studies in Conversation Analysis*, J. Maxwell Atkinson, John Heritage, eds., London, Cambridge University Press, pp. 225–246.
- Haddington, P., Keisanen, T., Mondada, L., Nevile, M. 2014.
 Multiactivity in Social Interaction: Beyond multitasking,
 Amsterdam/Philadelphia: Benjamins.
- Kendon, A. 1988. How gestures can become like words, in Poyatos (ed.) Cross-cultural perspectives in non verbal communication, Toronto: Hogrefe.
- Langacker, R. W. 1988. A view of linguistic semantics. In B. Rudzka-Ostyn (Ed.), *Topics in cognitive linguistics* (pp. 49–90). John Benjamins.
- Linell, P. (2009). Rethinking Language, Mind and World dialogically: Interactional And Contextual Theories Of Human Sense-Making. Charlotte, NC: Information Age Publishing.
- McNeill, D. 1992. Hand and Mind: What Gestures Reveal About Thought. Chicago: University of Chicago Press.
- Mondada, L. 2009. The Methodoldical Organization of Talking and Eating: Assessments in Dinner Conversations. Food Quality and Preference
- https://doi.org/10.1016/j.foodqual.2009.03.006 Mondada, L. 2016. Challenges of multimodality: Language and the
- body in social interaction. *Sociolinguistics*. Volume 20, issue 3: 336-366.
- Mondada, L. 2019. Contemporary issues in conversation analysis: Embodiment and materiality, multimodality and multisensoriality in social interaction. *Journal of Pragmatics*, 145, 47–62.
- Morgenstern, A. 2014. Children's multimodal language development. In Christiane Fäcke (Ed.). *Manual of language acquisition*. Berlin/Boston: De Gruyter, 123-142.
- Morgenstern, A. 2022. Children's multimodal language development from an interactional, usage-based, and cognitive perspective. WIREs, issue e1631.
- Morgenstern, Aliyah, and Goldin-Meadow, Susan (Eds.). 2021.

 Foreword Gesture as part of Language or partner to Language. In A. Morgenstern & S. Goldin-Meadow (Eds.), Gesture in language: Development across the lifespan (pp. 365–370). De Gruyter Mouton; American Psychological Association.

- Morgenstern A. & Parisse C. (Eds) 2017. *Le langage de l'enfant de l'éclosion à l'explosion*, Paris : Presses de la Sorbonne Nouvelle
- Morgenstern, A., Caët, S., Debras, C., Beaupoil-Hourdel, P, and Le Mené, M. 2021. Children's socialization to multi-party interactive practices: Who talks to whom about what in family dinners. In Letizia Caronia (ed.) Language and Social Interaction at Home and in School. Amsterdam: John Benjamins, 46-85.
- Ochs, E. 1988. Culture and Language Development: Language Acquisition and Language Socialization in a Samoan Village. Cambridge: Cambridge University Press.
- Ochs, E. 2012. Experiencing language. *Anthropological Theory*, 12 (2): 142-160.
- Ochs, E., Pontecorvo, C. & Fasulo, A. 1996. Socializing taste. *Ethnos*, 61, 7–46.
- Ochs, E. & Schieffelin, B. 1984. Language acquisition and socialization: Three developmental stories. in *Culture theory: Mind, self, and emotion*, ed. by R Shweder & R. LeVine. Cambridge: Cambridge University Press.
- Schieffelin, B. B. & Ochs, E. 1986. Language Socialization. *Annual Review of Anthropology* 15, 163-191.
- Slobin, D. 1987. Thinking for speaking. Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society, pp. 435-445
- Tomasello, M. 2003. Constructing a Language: A Usage-Based Theory of Language Acquisition. Harvard University Press, 2003.