



HAL
open science

Processus de focalisation multimodale lors des dîners familiaux français: une comparaison entre familles parlantes et familles signantes

Karine Martel, Christelle Dodane, Marion Blondel, Fanny Catteau, Corrado Bellifemine

► To cite this version:

Karine Martel, Christelle Dodane, Marion Blondel, Fanny Catteau, Corrado Bellifemine. Processus de focalisation multimodale lors des dîners familiaux français: une comparaison entre familles parlantes et familles signantes. *Faits de langues*, 2024, 53 (2), pp.111-142. <hal-04778853>

HAL Id: hal-04778853

<https://hal.science/hal-04778853v1>

Submitted on 28 Dec 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

Multimodal focalization processes during French family dinners: a comparison between speaking and signing families

Processus de focalisation multimodale lors des dîners familiaux français : une comparaison entre familles parlantes et familles signantes

Martel Karine¹, Dodane Christelle², Blondel Marion³, Catteau Fanny^{3,4}, Bellifemine Corrado²

GRHAPES (INSEI)¹, CLESTHIA (Sorbonne Nouvelle)², SFL (CNRS-Paris8)³, Université de Poitiers⁴

RESUME :

L'organisation de son discours permet à la personne qui s'exprime de faire passer au premier plan une information nouvelle afin d'attirer l'attention de son interlocuteur. Ce processus de focalisation a été décrit dans les langues parlées et dans les langues des signes, mais rarement dans un contexte naturel d'interaction familiale. Dans cet article, nous étudions la focalisation dans sa dimension prosodique, en comparant les phénomènes et paramètres mobilisés dans les deux types de langues, dans un contexte privilégié pour observer la focalisation à l'œuvre dans les interactions parents-enfants : le dîner familial. Nous dressons l'inventaire des marqueurs prosodiques de focalisation mobilisés dans les deux langues et les deux modalités (orale et gestuelle), dans un échantillon de repas en LSF et en français (ANR DinLang). Nous observons que des caractéristiques prosodiques de la focalisation peuvent être formulées en termes inter-modaux et que la focalisation est marquée par la combinaison de phénomènes plus ou moins nombreux, qui apparaissent par contraste avec le contexte prosodique. Nous donnons l'exemple d'un patron qui nous semble intéressant à examiner dans le cadre de la focalisation, celui de la scansion, et soulignons l'intérêt de prendre en considération les langues des signes et les langues 'parlées' comme des langues incarnées.

ABSTRACT :

Speech planning allows the speaker to highlight new information in order to draw the interlocutor's attention. This focusing process has been described in spoken and sign languages, but rarely in a natural context of family interaction. In this article, we study focalization at its prosodic level, by comparing the phenomena

and parameters mobilized in the two types of language, within a privileged context, in order to observe focalization during parent-children interaction: the family dinner. We make an inventory of prosodic focalization markers, mobilized in both languages and both modalities (spoken and gestural), from a sample of videorecorded meals in LSF and spoken French (ANR DinLang). We observe that prosodic features of focalization can be established in inter-modal terms and that focus is marked by the combination of various phenomena, which appear in contrast with the prosodic context. We give the example of a pattern that seems interesting to us to examine in the context of focalization, the one of scansion, and we emphasize the interest of considering sign languages and ‘spoken’ languages as embodied languages.

Mots clés : focalisation, interaction, geste, prosodie

Keywords : focalization, interaction, gesture, prosody

1. Introduction

Le domaine de la focalisation est assez difficile à appréhender car les contours de la notion varient en fonction des disciplines (Cappeau et Hanote, 2012). De façon générale, on peut dire qu’il s’agit d’un procédé permettant de mettre en contraste ou en relief certaines parties du discours (syllabes, mots, syntagmes) au sein d’un énoncé, dans le but d’attirer l’attention de l’interlocuteur sur ces éléments (Groussier et Rivière, 1996 : 84). D’un point de vue formel, il s’agit donc de mettre en avant des éléments par rapport à d’autres éléments concomitants ou plus distants, par exemple dans le cadre de l’opposition information donnée / information nouvelle (dans ce cas, on parlera de focus informatif), mais plus généralement sur le plan de l’accessibilité des informations et de la manière dont elles sont considérées comme plus ou moins pertinentes dans le discours. Les moyens linguistiques de la focalisation relèvent aussi bien de la syntaxe (comme dans les constructions clivées telles que « *c’est moi qui l’ai cuisiné !* », cf. Palancar et Vanhove, 2021), que de la prosodie, et nous proposons ici d’explorer la question en nous centrant sur le niveau prosodique.

Dans les langues audio-vocales, sur le plan prosodique, la saillance représente l’un des moyens par lesquels se manifeste la focalisation, l’élément « focalisé » étant mis en avant grâce à différents paramètres prosodiques (allongement, hauteur, intensité) liés au niveau acoustique à la durée, la fréquence fondamentale (désormais *fo*) ou encore l’amplitude (par exemple, lorsqu’un locuteur dit : « *Je suis très content* », en produisant un accent sur la syllabe « très » matérialisé par un allongement de durée, une hauteur plus élevée et une intensité plus forte). Cette mise en saillance (la focalisation prosodique) « revêt une fonction déictique qui consiste à pointer sur un élément quelconque du message » (Lacheret-Dujour et François, 2003 : 7) et contribue ainsi à « élargir l’environnement cognitif mutuel des protagonistes » (Di Cristo, 2000 : 31).

Les paramètres vocaux de la prosodie ne sont cependant pas les seuls à marquer la focalisation dans les langues parlées, les marqueurs visuo-gestuels tels que les mouvements des sourcils et de la tête (Ambrazaitis et House, 2022 ; Dohen et al., 2006) ou encore les gestes manuels (Esteve-Gibert et al., 2022 ; Ferré, 2018 ; Rorher et al., 2022 ; Roustan et Dohen, 2010) jouent également un rôle non négligeable dans la production et la perception de la focalisation.

À ce propos, il est intéressant de s'interroger sur les paramètres gestuels de la prosodie utilisés pour marquer la focalisation dans les langues essentiellement visuo-gestuelles, telles que la Langue des Signes Française (LSF), afin de voir si ces paramètres correspondent à ceux que l'on trouve dans les langues parlées.

Dans cet article, notre objectif est de mettre au jour les différents paramètres prosodiques permettant de marquer la focalisation, qu'ils soient vocaux, manuels ou non manuels tels que les mouvements du buste, de la tête, des articulateurs de la face¹. Nous adoptons une approche intermodale et multiparamétrique, comprenant aussi bien une description intra-langue des caractéristiques de la focalisation en français parlé et en LSF, qu'une description interlangue permettant de comparer ces deux langues. Nous nous interrogerons ainsi sur les moyens prosodiques que le français parlé et la LSF utilisent pour hiérarchiser certaines informations dans le discours (Cappeau et Hanote, 2012). Pour ce faire, nous étudierons la focalisation en contexte naturel, à partir d'un corpus de dîners familiaux soit en français parlé, une langue majoritairement audio-vocale (mais pour laquelle nous proposons de prendre en compte la multimodalité de son expression), soit en LSF, une langue visuo-gestuelle. Ces dîners, enregistrés dans le cadre du projet DinLang (Morgenstern et al., 2021, Parisse et al., 2022), rassemblent des parents et leurs enfants, en interaction naturelle. Ils constituent selon nous un contexte d'analyse particulièrement approprié, non seulement parce que spontané, mais aussi parce que chaque membre de la famille doit au cours du repas imposer / partager son point de vue, mettre en œuvre des stratégies pour prendre part à la discussion collective, et enfin donner des consignes ou transmettre une nouvelle information. La situation du dîner au domicile fournit par conséquent les conditions optimales pour observer le phénomène de focalisation prosodique en raison de la dimension écologique des interventions des partenaires sociaux, impossible à reproduire dans des conditions expérimentales. Aussi, il existe encore très peu de travaux qui ont comparé la prosodie des langues parlées à celles des langues des signes (Fenlon et al., 2019 ; Lombart, 2021a,b ; Tang et al., 2010) ; notre étude s'inscrit à la suite de ces recherches en ajoutant un contexte d'analyse original : l'étude de la focalisation prosodique au cours de dîners familiaux.

¹ Nous empruntons ces termes à Chloé Thomas (Thèse en cours, Université de Rouen) qui suggère d'employer « articulateurs de la face » plutôt que « expressions faciales » pour rester à un niveau de description formelle, comme pour les autres articulateurs, en amont de l'interprétation qu'implique « expression ».

Dans un premier temps, nous ferons un état de l'art sur les différents types de focalisation prosodique et leurs différents marqueurs multimodaux. Nous présenterons ensuite les deux dîners tirés du corpus DinLang (en cours de constitution), en français et en LSF, qui nous permettront de comparer la focalisation prosodique en langue dite "parlée" et en LSF. Les premiers résultats de nos analyses feront l'objet d'une discussion, en soulignant les points communs entre les deux dîners pour ce qui est de l'expression prosodique de la focalisation, les éventuelles spécificités et les pistes de généralisation associées à la comparaison interlangue et intermodale dans un même contexte dinatoire.

2. État de l'art : la focalisation dans tous ses états

Nous commencerons par présenter les différents types de focalisation et la façon dont ils sont marqués vocalement au niveau prosodique. Puis, nous nous intéresserons à la dimension gestuelle qui peut être associée à la prosodie vocale (pour les langues « parlées ») ou lui être substituée (pour les langues des signes). L'originalité de notre approche est de regrouper, dans cet état de l'art, la gestualité présente dans l'expression des sujets « parlants » et la gestualité intégrée dans un système linguistique essentiellement visuo-gestuel (celui d'une langue des signes), chez les sujets « signants ».

2.1. Le focus informatif dans la modalité *audio-vocale*

Nous avons vu plus haut que lorsqu'un locuteur parle, il met oralement en relief certains constituants de son énoncé. Ce procédé de focalisation est marqué au moyen de l'accentuation, l'intonation, la variation de débit et de registre tonal ; on parlera donc de **focalisation prosodique (ici vocale)**. La focalisation remplit plusieurs fonctions, soit de nature sémantico-pragmatique pour insister sur une information et la mettre en avant par rapport au reste du contenu (cf. focus informatif, contrastif / accent focal), soit de nature expressive pour transmettre ses émotions ou ses attitudes par rapport à un contexte donné et aux intentions du locuteur (cf. emphase / accent emphatique).

2.1.1. Focus prosodique large vs. focus prosodique étroit

La linéarité du langage conduit le locuteur à effectuer des choix dans sa façon d'agencer les différents termes de ses énoncés. C'est en fonction du degré de pertinence de ces informations, relatif aux connaissances partagées des partenaires de communication et au contexte, que le locuteur opère de tels choix (Wilson et Wharton, 2006). Certaines informations jouent un rôle secondaire, tandis que d'autres sont considérées comme primordiales pour une compréhension mutuelle au fil des échanges verbaux. La communication d'informations nouvelles, en l'occurrence, est souvent effective sur un arrière-plan propositionnel comprenant des

informations présumées déjà connues (activées sur le plan cognitif chez les interlocuteurs).

Selon l'approche informationnelle, l'analyse d'un énoncé implique trois niveaux : grammatical, sémantique et thématique. C'est de ce troisième niveau que relève l'étude du thème (ou *topic* correspondant à l'être ou la chose dont il est question et qui est souvent connu par le contexte) et celui du rhème (l'élément nouveau qui contient une information sur le thème). Le rhème, généralement présenté comme le complément du thème dans l'organisation de la structure énonciative, est aussi appelé commentaire ou propos. Rossi (1999) définit ce terme comme le « noyau d'énoncé », porteur de l'information de premier plan, susceptible d'agir sur l'interlocuteur et de modifier ses représentations. Il parle également de « rhème complexe » qui correspond à un rhème contenant un ou plusieurs constituants extra-posés par l'intonation sans dislocation syntaxique et susceptibles de recevoir l'accent de focalisation (Lacheret-Dujour et Beaugendre, 1999 : 271). Le focus informatif désigne donc le rhème, c'est-à-dire la nouvelle information exprimée par le locuteur à propos d'une information déjà connue de l'auditeur. La notion de rhématisation est souvent liée à la présence d'un accent nucléaire en anglais. Il existe deux types de focus informatif (Lambrecht, 1994 ; Godement-Berline, 2018) : le focus informatif large (FL), où l'énoncé entier constitue une nouvelle information (par exemple, dans la réponse à la question : « *Qu'est-ce qu'il a ?* » - « *Il vient juste d'apprendre qu'il avait raté son examen* », le focus informatif est large car aucune des informations contenues dans l'énoncé n'a déjà été formulée dans la question qui précède) et le focus informatif étroit (FE), où seule une partie de l'énoncé constitue une nouvelle information (par exemple, dans la réponse à la question : « *Où sont allés tes parents ?* » - « *Ils sont allés faire des courses* », le focus informatif est étroit, car seule une partie de l'information n'a pas déjà été formulée dans la question qui précède).

Dans les langues « parlées », ces deux types de focus informatif connaissent une réalisation prosodique vocale différente : dans le cas du FL, tous les constituants majeurs de l'énoncé reçoivent une accentuation alors que dans le cas du FE, seul un constituant reçoit une accentuation (Godement-Berline, 2018). Plus spécifiquement, le FL reçoit le phrasé par défaut en groupes accentuels et intonatifs et la syllabe nucléaire est attribuée à la syllabe accentuée correspondant à la dernière syllabe pleine en français (Ferré, 2019). Les constituants très courts sont parfois dépourvus d'accent et sont phrasés dans le même groupe que le constituant adjacent (Godement-Berline, 2018). Du point de vue de l'intonation, l'énoncé en FL est « souvent uniquement marqué par la présence d'un ton final descendant » (Féry, 2001 : 176), mais aussi par l'allongement de la syllabe finale. Dans certains contextes, l'accent nucléaire ne porte pas sur le dernier item lexical. C'est le cas lorsqu'un énoncé contient un « post-rhème » (Morel et Danon-Boileau, 1998), qui sera désaccentué (par exemple dans les marqueurs de discours finaux, de point de vue ou discursifs situés en position post-focale, ainsi que les dislocations à droite

et les appendices, Ferré, 2019). À propos du marquage prosodique du FE, on observe également un contour intonatif terminal sur le constituant focalisé et une désaccentuation (Delattre, 1966) sur la séquence qui suit ou précède, mais celle-ci n'est pas systématique (Godement-Berline, 2018). Dans Beyssade et al. (2009), le contour terminal ne tombe sur le constituant focalisé que dans 60% des cas et dans 40% des cas, il tombe à la fin de l'énoncé. Le constituant focalisé peut porter un accent initial, qui forme un arc accentuel avec l'accent final du constituant (Di Cristo, 1999b). En FE, la règle d'accentuation du dernier item lexical ne s'applique pas et n'importe quelle syllabe de l'énoncé est susceptible de porter l'accent de focalisation dont le trait acoustique essentiel est une montée importante de la hauteur de la voix (Rossi, 2011 : 125), y compris des syllabes de mots grammaticaux monosyllabiques. Le FE peut avoir plusieurs fonctions, qui sont décrites ci-dessous.

2.1.2. Le focus contrastif

La plupart des travaux sur le FE sont basés sur l'analyse d'un certain type de focus prosodique, le focus contrastif (exemple : « *C'est toi qui a donné le livre à Marie ? Non c'est Paul qui le lui a donné* »), (Ferré, 2019). Le focus contrastif porte non pas sur l'action de donner le livre à quelqu'un (qui fait partie des connaissances partagées et qui est en dehors du domaine focal), mais sur la personne qui a donné le livre. Ici, le constituant focalisé entre dans une relation de contraste avec une information qui précède. Dans ce cas, on peut considérer qu'il s'agit de marquage de focus informatif étroit. Godement-Berline (2018) affine la description de l'accent contrastif en reprenant la notion de contraste de Vallduvi et Vilkuna (1998). Selon ces derniers, une information est contrastive dès que sa formulation génère un ensemble d'alternatives dont elle fait partie (par exemple « *Je suppose que tu as mal à la tête* » > « *Non, j'ai mal au ventre* »).

Il existe deux types de marquage du focus contrastif. Dans le premier cas, le constituant focalisé est suivi de constituants au statut informationnel ancien et l'on observe une relation similaire à celle observée pour le marquage du focus informatif étroit : le contour focalisé porte un contour intonatif terminal réalisé avec un ton haut porté par la syllabe initiale ou finale du constituant focalisé associé à une augmentation de durée et la séquence qui précède ou qui suit est désaccentuée (Di Cristo, 1999b ; Jun et Fougeron, 2000 ; Dohen et Loevenbruck, 2004). Selon Di Cristo (1999b), le constituant focalisé est marqué par un accent initial rythmique ou emphatique. Dans le second cas, qualifié par Lambrecht de « topic-contrastif », Beyssade et al. (2004) observent la présence d'un accent initial spécifique sur le constituant focalisé portant sur la première syllabe et réalisé avec une forte augmentation de la hauteur, un allongement du segment initial de la syllabe accentuée et une augmentation de l'intensité. Le focus contrastif est en effet marqué par un accent qui est perceptivement plus saillant que l'accent par défaut dans les énoncés en FL (Katz et Selkirk, 2011 ; Kramer et Swerts, 2001). Cet accent contrastif est marqué

par un mouvement mélodique plus large, une intensité plus élevée et/ou un allongement syllabique (Katz et Selkirk, 2012 : 81). Il est réalisé à partir d'un mouvement intonatif montant-descendant dont le sommet est aligné avec la première syllabe et un allongement significatif de la consonne d'attaque (Astésano et al., 2004). Précisons que si le focus contrastif se manifeste à travers le FE, il ne correspond néanmoins pas systématiquement au focus informatif tel que nous l'avons présenté comme étant l'élément nouveau/non connu de l'énoncé. Le focus contrastif peut aussi bien concerner le thème que le rhème du fait que les termes mis en relief à l'intérieur du thème ou du rhème sont en relation de contraste avec d'autres référents dans le discours ou le contexte situationnel et toujours relié au dynamisme communicatif (Lambrecht, 1994 ; Stavropoulou et Spiliotopoulos, 2011).

2.1.3 Le focus étroit emphatique

Pour Di Cristo (1999a), l'emphase désigne l'ensemble des **fonctions expressives de la focalisation prosodique**. Les caractéristiques prosodiques ont été notées pour l'**emphase** dans les **énoncés non contrastifs**, comme l'illustre l'exemple suivant : « *C'est vraiment magnifique !* ». Dans cet exemple, la première syllabe du dernier mot porte un accent emphatique également appelé accent d'insistance, qui a pour effet de renforcer la qualité « magnifique » présumée et qui tombe en dehors du focus. Contrairement au focus contrastif, selon Bagou (2001 : 47), « l'implication emphatique est envisagée comme un élément de communication de nature émotive, planifiée et stratégique, permettant au locuteur de signifier la présence d'événements particulièrement essentiels à la compréhension de son message par le destinataire ». On trouve généralement deux types de fonctions emphatiques, la fonction d'insistance associée à un accent initial de mot et utilisée pour souligner à l'interlocuteur l'importance d'une partie ou de la totalité du contenu de l'énoncé (De Paolis et al., 2022), et la fonction expressive correspondant à l'expression des émotions et des attitudes (Di Cristo, 1999a). Si les fonctions sont différentes entre les divers types de focus étroits, le marquage n'est pas nécessairement très différent sur le plan prosodique. L'accent initial d'insistance est également marqué par une augmentation de hauteur, de durée et d'intensité (Di Cristo, 1999b). Cependant, l'allongement de durée concerne surtout la consonne initiale et selon Séguinot (1977), la hauteur joue un rôle plus important que les deux autres paramètres. Lorsque la syllabe commence par une voyelle, un coup de glotte est généralement réalisé avant l'accent (Di Cristo, 1999b). Dans le second type (l'accent expressif), les attitudes sont globalement marquées par une forte intensité des syllabes accentuées, ainsi que par des groupes prosodiques courts et isochroniques. Comme pour le focus contrastif, les accents emphatiques se caractérisent par une augmentation de hauteur et d'intensité nettement supérieure aux accents des unités voisines. En français, on remarque une forte élévation de hauteur et de durée sur la

syllabe qui porte le contour terminal de l'énoncé, ce qui a pour effet d'augmenter l'amplitude de ce contour (Di Cristo, 1999b).

Nous allons maintenant porter notre attention sur le marquage gestuel de la focalisation. Certains chercheurs s'intéressent aux ressources prosodiques vocales et gestuelles en comparant les langues parlées et les langues des signes (Lombart, 2021b), et ce, notamment, dans le cadre du focus contrastif, présenté plus haut.

2.2. Le marquage de la focalisation prosodique au niveau gestuel

Dans ce travail, la focalisation est abordée aussi bien sur le plan vocal que gestuel et il s'agit de proposer un cadre de codage inédit permettant de comparer la gestualité chez les sujets « parlants » à celle des sujets « signants ». La partie qui suit expose le marquage de la focalisation au niveau gestuel.

2.2.1 La focalisation dans les langues des signes

L'étude de la prosodie des langues des signes (LS) est bien plus récente que celle des langues parlées, c'est pourquoi il convient de donner quelques repères généraux sur le domaine avant d'aborder la prosodie de la focalisation dans ces langues gestuelles. Les premiers travaux diffusés en prosodie des LS ont concerné le débit, les pauses et tenues dans le flux signé, la durée des syllabes en langue des signes américaine (American Sign Language : ASL, Grosjean et Lane, 1977 ; Wilbur et Nolen, 1986, entre autres). Peu à peu, des études se sont intéressées à la structuration prosodique des énoncés, voire de séquences discursives (Boyes-Braem, 1999) et progressivement, c'est l'ensemble des marqueurs manuels et non manuels qui a été pris en compte pour décrire les phénomènes prosodiques. L'enjeu des outils de mesure du signal gestuel a été introduit avec l'apparition des nouvelles technologies, comme la capture de mouvement. En parallèle de ces avancées dans la description articulatoire ou "phonétique" des LS, la dimension interlinguistique a pris de la consistance avec une variété plus grande de LS étudiées (suisse allemande ou *DSGS*, israélienne ou *ISL*, japonaise ou *JSL*, néerlandaise ou *NGT* et *LSF*, entre autres). Les principales questions étudiées à ce jour, en prosodie des LS, rejoignent les études linguistiques des LS en général : Quels sont les points communs et différences entre prosodie des langues parlées et prosodie des langues des signes ? Les différences liées au canal emprunté ont-elles un impact sur la réalisation ou sur la structure prosodique des LS ?

La prosodie des LS ne passe en effet pas par la voix mais par le reste du corps et par les mouvements produits par les articulateurs multiples : doigts, mains, bras, épaules, tête, sourcils, lèvres, etc. Une part des travaux, inscrits dans une approche prosodique phonologique (Nespor et Vogel, 1986) et appliquée aux LS, tend à distinguer les articulateurs (brachio-) manuels et non manuels (sourcils, tête, buste...), en les associant respectivement aux aspects rythmiques et aux aspects intonatifs de la

prosodie (Navarrete-González, 2021, Sandler, 1999 ; Sandler et Lillo-Martin, 2005). D'autres travaux nuancent cette dichotomie et adoptent une approche descriptive articulatoire qui place le *mouvement* au centre de l'étude prosodique (Wilbur, 1999 ; Wilbur et Malaia, 2018 pour l'ASL ; Boyes-Braem, 1999 pour la DSGS ; Blondel et Le Gac, 2006 et Catteau, 2020 pour la LSF, entre autres). Pour ces auteur·es, la durée, l'amplitude et la vitesse du mouvement sont trois paramètres prosodiques des LS, qui correspondent peu ou prou aux paramètres de durée, intensité et fréquence fondamentale des langues parlées.

Catteau (2020) et Blondel (2020) proposent que la dichotomie rythmique / intonative n'a pas lieu d'être dans la mesure où chaque articulateur mobilisé est susceptible d'être modulé selon ces trois paramètres, même si les mouvements sont parfois d'amplitude si réduite qu'on les associe facilement à une position finale plutôt qu'au mouvement en tant que tel (par exemple, une *position haute* des sourcils plutôt qu'un *haussement* de sourcils). La réalisation d'un phénomène prosodique, comme la mise en saillance par exemple, se caractérise par l'augmentation relative de l'amplitude non seulement du mouvement des mains, mais également du mouvement des sourcils ou de la tête (comme nous le détaillons plus loin). En d'autres termes, nous proposons que c'est bien la *combinaison* des différents articulateurs mobilisés dans l'expression des LS et des paramètres prosodiques qui leur sont associés qui permet de former les contrastes et les phénomènes prosodiques qui nous intéressent, dont celui de la focalisation.

Les rares travaux disponibles sur le sujet décrivent les moyens (prosodiques) d'expression de la focalisation et visent en priorité à 'interroger' l'éventuelle spécificité des langues liée à leur modalité visuo-gestuelle. Ces premiers travaux se situent majoritairement à l'interface prosodie / syntaxe, et nous indiquons ci-après les éléments d'ordre syntaxique qui ont une répercussion sur les contours prosodiques. Ainsi, Wilbur (1994a, b, 1997, 1999) pour la langue des signes américaine, Nespor et Sandler (1999) pour la langue des signes israélienne observent que le focus non contrastif se place en position finale d'énoncé. Wilbur (1997) précise qu'il existe des 'marqueurs de focalisation' (*focusers*) lexicaux et que la position finale correspond à une position saillante. Dans une étude de la prosodie en langue des signes de Suisse allemande, Boyes-Braem (1999) remarque également que, pour créer un effet d'insistance sur un élément, on a tendance à déplacer un signe en position finale, et, éventuellement à le ré-dupliquer. Ces phénomènes de réduplication dans l'expression du focus contrastif sont d'ailleurs illustrés par Blondel (2003), dans un corpus longitudinal d'interactions familiales en LSF. Wilbur (1994b) mentionne également les 'clivées avec mots interrogatifs' (*wh-clifs*), ou 'pseudo-clivées', aussi appelées *fausses questions* ou *questions rhétoriques* (Moody, 1983 : 94-95). Dans ces constructions, l'information de second plan est incluse dans un syntagme interrogatif et les constituants focalisés passent au premier plan en position finale. En complément de cette interface syntaxe / prosodie et en ce qui concerne les marqueurs

prosodiques manuels, selon Sandler (1999), la réduplication d'un signe, sa tenue ou le fait qu'il soit suivi d'une pause sont trois éléments pour rendre saillant un élément en fin d'énoncé en langue des signes israélienne. Remarquons que les trois phénomènes aboutissent à une forme d'allongement de la durée. La durée du signe, la distance parcourue et la vitesse sont aussi mises en rapport dans les travaux de Wilbur. Blondel (2003) donne des exemples pour illustrer ces phénomènes en LSF avec des énoncés dans lesquels des unités lexicales sont suivies d'une légère pause, ou articulées très distinctement (amplitude forte, rythme lent). Or, l'ensemble de l'information portée par les signes en question est nouvelle, même si on peut considérer qu'il y a des degrés de saillance variés à l'intérieur de la proposition, avec des constituants qui le sont plus que d'autres.

En correspondance avec la variation de *f₀* en langue parlée, Wilbur (1999) relève un 'pic de vitesse' (*high peak of velocity*). La mesure du 'déplacement' (*displacement*) résulte de la contrainte d'allongement de la durée en finale et de la recherche de la proéminence sous forme de pic de vitesse. Ainsi, si deux signes sont en position finale avec un allongement de la durée approprié à la position finale, pour que le signe A soit perçu comme plus saillant que B, il devra avoir un pic de vitesse plus important et donc parcourir une plus grande distance pour la même durée.

D'autres voies de prospection sont ouvertes à partir des travaux sur l'accent. Notons parmi elles l'observation selon laquelle un signe accentué est signé plus haut (dans l'espace) que le signe correspondant non accentué. Cette remarque est faite notamment dans le cadre des travaux sur la focalisation en langue des signes néerlandaise (van der Kooij et al. 2006).

A propos des articulateurs de la face et du regard, van der Kooij et al. (2006) relèvent un contact du regard maintenu pendant l'articulation des constituants focalisés. Blondel (2003) mentionne également le rôle du haussement de sourcils pour marquer un focus contrastif en LSF.

Au sujet des mouvements du buste et de la tête accompagnant les éléments focalisés, Wilbur (1994b : 669) rapporte le hochement simple de tête (*single head nod*), le mouvement de tête vers l'arrière (*head back*), le fait de rentrer le menton (*chin tucked*) et ne clôt pas la liste. Wilbur et Patschke (1998) ont remarqué que le fait de pencher le buste vers l'avant ou l'arrière peut indiquer une emphase sur des items lexicaux, sur des catégories d'inclusion (*even*, 'même') et d'exclusion (*only* 'seulement') et des focus contrastifs. Van der Kooij et al. (2006) relèvent à partir de leurs données en langue des signes néerlandaise que ce sont des mouvements latéraux du buste qui sont impliqués dans les focus contrastifs.

Pour les LS, ces marqueurs gestuels (qu'ils soient manuels ou non manuels) constituent en tant que tels, et à eux seuls (i.e. sans vocalisation), les indicateurs prosodiques de l'organisation de l'information dans le flux signé. Comme les langues parlées, jusqu'à récemment, sont souvent considérées à travers le flux vocalisé (pour cette même dimension prosodique), la mobilisation du reste du corps est restée peu explorée à l'exception des travaux sur l'association gestes-voix (Dohen et

Loevenbruck, 2009 ; Dohen et al., 2011 ; Ferré, 2019 ; Brown et Prieto, 2021 entre autres). Nous proposons de revisiter ces éléments gestuels dans les langues parlées à la lumière des connaissances établies pour la prosodie gestuelle des LS.

2.2.2 La focalisation gestuelle dans les langues dites « parlées »

Différents articulateurs sont pris en considération dans les études concernant la gestualité dans les langues parlées : les articulateurs de la face, la tête, le buste et les mains. Ces gestes manuels et non manuels sont étudiés à un niveau prosodique de façon similaire aux marqueurs vocaux et plusieurs études rendent compte de la synchronisation et désynchronisation entre les deux types de marqueurs (Graf et al., 2002 ; Granström et House, 2005 ; Loehr, 2014 ; Lombart, 2021).

En ce qui concerne la perception de la focalisation, House et al. (2001) relèvent que les mouvements des sourcils et de la tête sont de puissants repères visuels des syllabes proéminentes au niveau prosodique et que la sensibilité perceptuelle à la synchronisation est de l'ordre d'une durée de syllabe typique de 100 à 200 ms. Par ailleurs, les froncements de sourcils sont notamment associés au marquage d'un focus, quel qu'en soit le type (Krahmer et al., 2002). En particulier, il a été montré une étroite relation entre les haussements des sourcils et la mise en saillance prosodique vocale (Granström et al., 1999a, 1999b). En outre, Dohen et Loevenbruck (2005) regroupent différentes manifestations 'gestuelles faciales' du focus contrastif sous le terme d'*hyper-articulation* (augmentation de l'amplitude des mouvements d'ouverture et de fermeture des lèvres et/ou de la mâchoire, accroissement de l'étirement ou resserrement des lèvres), suggérant ainsi l'implication de ces parties du visage dans le marquage du focus.

En utilisant la capture de mouvement, Roustan et Dohen (2010) ont étudié l'articulation entre le focus contrastif vocal et différents types de gestes manuels à production induite (pointages, gestes de battement, mouvements liés au contrôle moteur d'objets) auprès de dix locuteurs lors d'une tâche d'élicitation de focus à partir de quatre stimuli sonores distincts constitués d'énoncés SVO avec des schémas syllabiques identiques pour chaque mot. Les auteurs ont observé un alignement temporel entre l'apex du geste - indépendamment de sa nature - et la syllabe du mot portant le focus au niveau prosodique. De la même manière, Ferré (2018) a mené une étude perceptive chez quinze locuteurs qui devaient identifier la présence d'emphase au niveau prosodique dans dix énoncés préalablement enregistrés (cinq neutres et cinq avec emphase) et accompagnés de différents types de gestes (pointages, battements). Les résultats montrent que, si l'identification des énoncés avec emphase était réussie, les locuteurs jugeaient également les énoncés neutres comme emphatiques lorsqu'un geste de battement était produit (mais pas en présence de pointages ou en l'absence de gestes). De plus, les gestes produits avec une plus grande amplitude menaient les locuteurs à identifier plus fréquemment de

l'emphase aussi dans les énoncés neutres. Au vu du caractère hautement contrôlé de ces observations, on pourrait se demander si les mêmes relations multimodales de la focalisation se manifestent dans des contextes écologiques issus de données naturelles et spontanées.

En outre, lorsque l'on prend en compte les mouvements du buste, on s'intéresse notamment aux changements de posture. Certains auteurs considèrent que le corps et sa posture participent activement à l'échange et établissent des relations socio-interactives (Mondada, 2017 : 185). Les études qui analysent de façon systématique la relation entre les changements de posture et la focalisation sont plus rares. Par exemple, Ronveaux et Simon (2005) observent comment les variations intonatives au niveau prosodique, les déplacements du corps et les changements de posture contribuent de façon multimodale à la gestion (en particulier à l'ouverture et à la clôture) d'activités effectuées par l'enseignant et les élèves à l'école. Nous en savons beaucoup moins sur les interactions en famille.

La (co-)construction du discours sur les différents plans linguistiques, se fait donc non seulement par le verbal-vocal mais aussi grâce à l'utilisation de plusieurs *Gestalts* (Mondada, 2017 : 185), autrement dit différentes ressources plurisémiotiques étroitement imbriquées pour garantir la richesse de la transmission du message. Il est donc crucial de travailler sur la coordination et l'éventuelle synchronisation entre les niveaux prosodique-vocal et gestuel dans le domaine de la focalisation.

Nos questions de recherche

À l'aune de cet état de l'art, nous proposons de confronter à des données d'interactions spontanées, recueillies en contexte familial, les hypothèses et questions suivantes. Premièrement, il semble difficile d'établir un 'profil type' de l'aspect formel de l'élément focalisé, que ce soit dans les langues parlées ou dans les langues des signes étudiées, car, comme le soulignent Crasborn et van der Kooij (2013), les moyens mis en œuvre pour la LS étudiée sont non seulement contraints par la structure phonologique d'un signe, mais également par la dimension pragmatique de l'interaction. Le plus souvent, on observe un effet produit par un ensemble de facteurs plutôt que par un seul (par exemple, le pointé est accompagné d'un mouvement de tête assez prononcé et de labialisation, le mouvement manuel est ample et tendu). En outre, il est parfois difficile de déterminer quelle manifestation est liée à quelle cause (quel type d'emphase par exemple).

Deuxièmement, peut-on considérer qu'il existe, de manière assez systématique, des équivalents *intermodaux* pour marquer la focalisation d'une information en LSF et en français 'parlé' (par exemple, réaliser un signe plus haut que les autres – quand cela ne répond pas à une contrainte de référence spatiale de type grammatical – serait-il équivalent à produire un pic de hauteur dans une courbe mélodique sonore) ? Lié à cela, quelles sont les spécificités éventuelles dues à la prépondérance d'une modalité (visuo-gestuelle / audio-vocale) ? Et doit-on nuancer le caractère spécifique selon qu'il est attribué aux langues ou à leur(s) modalité(s) ?

Enfin, en nous appuyant sur l'analyse des données DinLang (Morgenstern et al., 2021) ainsi que sur nos précédentes études, nous accorderons une attention particulière à l'émergence de patrons contrastifs, autrement dit de phénomènes formels et dynamiques qui rendent l'information saillante en contraste avec le contexte, ainsi qu'à la distribution du focus.

3. Méthodologie pour les dîners en famille

Le projet ANR DinLang² a plusieurs objectifs : étudier les liens entre formes, pratiques multimodales du langage et pratiques dînatoires dans les situations de repas familiaux, améliorer notre connaissance sur l'articulation des activités langagières et dînatoires, et la transmission des savoirs culturels, pratiques et langagiers au sein de la famille, étudier le développement du langage *in situ*, dans sa multimodalité et dans la dynamique des interactions.

3.1. Le dîner, un cadre privilégié pour étudier la focalisation

Le dîner est un contexte propice à l'interaction spontanée entre adultes et enfants, mais également entre adultes et entre enfants. C'est aussi un contexte dans lequel des informations sont échangées sur la journée écoulée et sur les projets des uns et des autres. Les membres de la famille ayant été séparés pendant la journée, tout en ayant un certain degré de connaissance partagée, leurs échanges oscillent entre ce qui relève du proximal spatio-temporel (l'ici et maintenant du repas) et le distal (l'ailleurs dans le temps et dans l'espace). Parmi la somme des informations échangées, on s'attend à observer des stratégies pour faire ressortir celles qui sont les plus significatives aux yeux du participant qui s'exprime. Par ailleurs, il s'agit d'un contexte où chacun doit trouver, négocier sa place dans l'interaction, et éventuellement la maintenir. Le repas quotidien en famille répond par ailleurs à des contraintes logistiques et pragmatiques : les parents s'assurent que chacun est servi, se nourrit et 'se tient bien à table', chacun (selon son âge et ses attributions) remplit des tâches nécessaires au déroulement du repas comme servir, desservir, et doit coordonner les actions du repas aux pratiques langagières. L'aspect routinier du repas quotidien permet à la famille d'oublier petit à petit le dispositif d'enregistrement et de retrouver le spontané des interactions *in situ*.

Comme la population ciblée inclut des fratries d'au moins deux enfants, dont l'un est d'âge scolaire (3-10 ans), les données DinLang permettent de considérer différents types de discours, dont le langage adressé ou non à l'enfant, sous toutes leurs formes et fonctions (informatif, explicatif, narratif, injonctif, interrogatif...), et ouvrent des pistes pour considérer le

² <https://anr.fr/Projet-ANR-20-CE27-0010>

développement et la maîtrise progressive de certaines compétences et stratégies interactionnelles avec l'âge (Morgenstern et al., 2021).

3.2 Le dispositif et les données sélectionnées

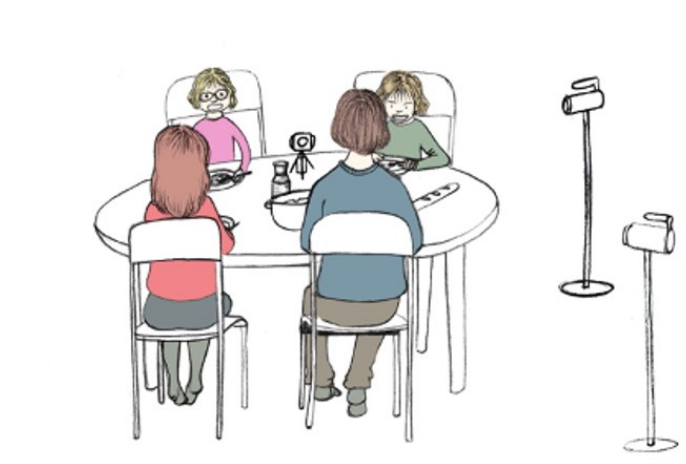


figure 1 : Disposition des trois caméras au cours des enregistrements des dîners familiaux

Le recueil du corpus DinLang comprend une section pilote composée de données issues des projets Signes en Famille (Morgenstern et Caët, 2021) et Dîners familiaux (Morgenstern et al., 2021). Une partie du dispositif est commune à ces trois projets, à savoir l'utilisation de deux caméras classiques dotées de micros externes de qualité et disposées de façon à permettre une prise de vue depuis chacun des côtés de la scène du repas ; à partir des enseignements du corpus pilote, une caméra 360° a été ajoutée au centre de la table et offre une vue de face de chaque participant (Figure 1), ainsi qu'un enregistreur sonore à 360° posé à côté des dîners où les participants s'expriment vocalement.

Les données sur lesquelles nous appuyons cette première étude de la focalisation correspondent à deux dîners d'une trentaine de minutes chacun. L'un se déroule en français (corpus pilote), dans une famille avec trois garçons âgés de 2, 3 et 9 ans, l'autre dîner se déroule en LSF (corpus signes en famille) dans une famille avec deux filles de 6 et 2 ans et demi.

3.3. Relevé et notation des occurrences de focalisation

3.3.1. Les étapes du protocole

Le relevé des items focalisés s'est fait dans un premier temps³ de manière individuelle, deux d'entre nous identifiant les occurrences de focus (large ou étroit) dans le dîner en français, et les deux autres chercheuses identifiant les occurrences en LSF. Pour le français parlé, nous avons utilisé les mesures fournies par le logiciel Praat (Boersma et al., 2023) et effectué un accord entre codeurs. Pour la LSF, le premier relevé s'est effectué sous Elan (Sloetjes et Wittenburg, 2008), en ajoutant des pistes *ad hoc* au schéma d'annotation commun (aux différents sous-groupes thématiques du projet DinLang, cf. Parisse et al., 2022). Là encore, un premier accord inter-codeuses s'est fait en confrontant les items relevés au sein du binôme (chacune ayant annoté sur une piste séparée et aveugle aux annotations de l'autre codeuse).

Si les codeuses du français se sont appuyées principalement sur le repérage de l'accent de focalisation vocale dans un premier temps, à partir de l'écoute du signal sonore et dans un second temps, à partir d'analyses acoustiques réalisées avec le logiciel Praat, les codeuses de la LSF ne peuvent pas, à ce stade, exploiter encore le signal gestuel avec un outil de mesure équivalent (nous reviendrons sur cette question plus loin). Nous nous sommes donc appuyées sur la notion la plus englobante possible de la focalisation (voir en 2.1), soit notre perception d'une information (nouvelle, la plupart du temps, mais pas exclusivement) sur laquelle le locuteur-signeur souhaite attirer l'attention en le plaçant au premier plan de la structure informationnelle, et nous avons commencé à inventorier les aspects formels, ou marqueurs potentiels, associés à ces balises-focus.

Notre analyse porte sur les vingt premières minutes de chacun des dîners (en français et en LSF). Bien que nous ne nous soyons pas appuyés sur des analyses quantitatives systématiques dans nos données, que nous envisageons d'explorer plus finement dans l'avenir, nos observations préliminaires nous ont néanmoins permis d'identifier des régularités au sein d'une même langue, entre les langues et entre les modalités (4.1).

3.3.2 Paramètres analysés et réflexions sur le *template*

En ce qui concerne le français parlé, nous avons codé le focus prosodique, indiqué s'il était large ou étroit, le contexte de la focalisation, le type de focalisation (informatif ou emphatique) si l'accent de focalisation était réalisé au niveau vocal et/ou gestuel, la modalité de l'énoncé (assertion, injonction, question, interpellation, etc.), la durée totale de l'énoncé (en millisecondes), l'intensité moyenne de l'énoncé (en dB), le marquage acoustique des différentes proéminences (*fo* en Hz,

³ Nous faisons depuis évoluer notre protocole de sélection et d'annotation des séquences focalisées en homogénéisant l'entrée dans les deux types de données (FR/LSF) en nous appuyant sur un template partagé sous ELAN. Nous espérons ainsi éviter un risque de circularité qui consisterait à partir d'un marqueur spécifique pour obtenir les autres marqueurs potentiels.

intensité relative en calculant le rapport d'énergie RMS (root-mean square en dB) de la totalité de l'énoncé étudié par rapport au pic d'intensité localisé sur la syllabe correspondant à l'accent de focalisation (en adaptant légèrement la formule de Shi et al., 1998) et durée des différentes syllabes (en ms), la configuration du contour d'intonation (M pour montant, D pour descendant, P pour plat, etc.), le regard (le point d'attention du locuteur), les mouvements des mains, de la tête et du corps, si ces mouvements concernent un seul articulateur ou plusieurs. Les mouvements ont été codés par un expert de la gestualité.

Parmi les éléments encodés pour le français parlé, certains ont trouvé leur équivalent en LSF : le locuteur/sgneur, la transcription de l'énoncé qui contient le focus, le *timecode* du début d'énoncé, sa modalité énonciative. Dans un premier temps, nous avons compilé en une annotation l'ensemble des marqueurs prosodiques potentiels du focus en LSF (avec des exemples de codage en (1)), et ajouté une rubrique concernant l'intercodage pour vérifier les sources d'accord et de désaccord entre les deux codeuses.

(1) a. [*apres-APPEL_accent:acceleration-MD_hold*]

...indique *que nous* percevons une **accélération** sur la **main dominante**, que le signe focalisé vient après un **geste d'appel** et semble **tenu** ;

b. [*synt:repet_accent:acceleration-amplitude-trans-MD(1)
accent:acceleration-trans-MD(2)_accent:acceleration-trans-MD(3)
_accent:acceleration-trans-MD(4)*]

...indique que les segments focalisés (1, 2, 3 ou 4) sont marqués par une **répétition**, une **accélération** ou une **amplitude** remarquable, un mouvement de **transition** remarquable

Dans un second temps, nous avons conçu un *template* « prosodie » sous Elan (Figure 2) pour relever de manière plus systématique les marqueurs prosodiques gestuels qui ont retenu notre attention lors de la première phase de notation. Ce *template* est conçu comme un module que l'on peut intégrer au *template* général commun aux différents domaines explorés dans le projet DinLang. Il comprend donc le même type de codage que le tronc commun pour indiquer quel participant s'exprime (M pour Mère, P pour Père, Ea pour l'enfant le plus âgé, Eb, pour le deuxième...). Les pistes dépendantes permettent de baliser les énoncés avec focus, d'y localiser précisément sur la ligne de temps l'élément focalisé, les phénomènes et marqueurs associés.

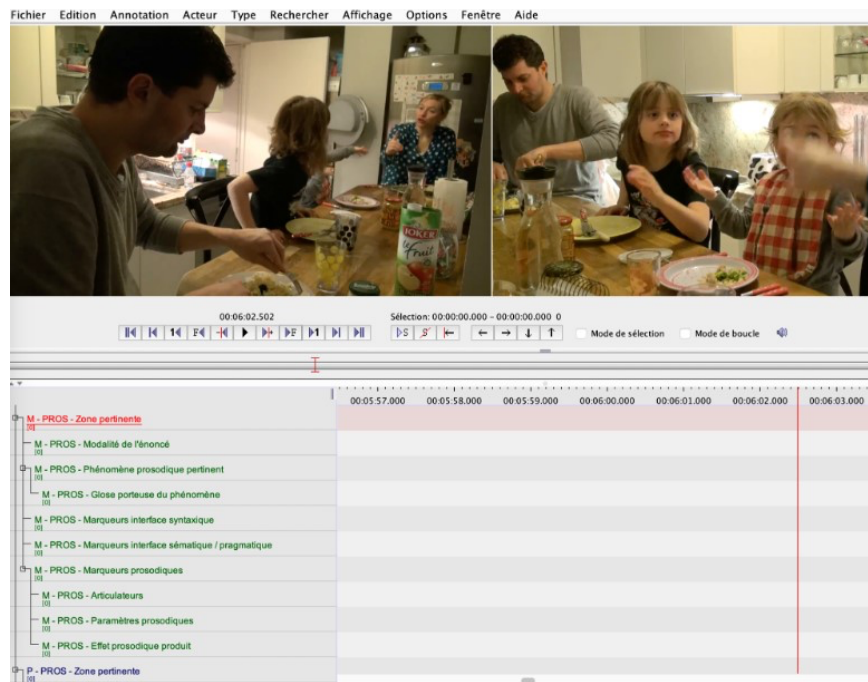


figure 2 : Sous-template « prosodie » sous Elan

4. Analyses

En français parlé nous avons annoté 68 occurrences de focalisation prosodiques à partir d'un extrait d'une vingtaine de minutes du repas sélectionné. Ces occurrences correspondent à des phénomènes repérés à partir des paramètres prosodiques vocaux, il s'agit d'un premier panel. Les interlocuteurs en ont réalisé davantage mais il a fallu écarter les productions non analysables en raison des chevauchements de parole et des bruits parasites liés à l'activité dinatoire (tintements des couverts dans les assiettes, les plats, etc.). Par ailleurs, il faut prendre en compte le fait que les personnes sont en interaction mais sont aussi en train de manger, ce qui réduit les temps de parole en comparaison avec une discussion hors repas. Les adultes formulent de nombreuses injonctions, répétitions pour faire en sorte que les enfants se nourrissent et ces derniers acquiescent souvent par un hochement de tête ou silencieusement. En LSF, la durée de l'échantillon d'annotation est plus courte, 13 minutes, mais comprend davantage d'occurrences (en proportion de la durée et en comparaison de l'échantillon en français) dès lors que le premier filtre pour la LSF s'appuie sur notre

perception d'un élément d'information focalisé, et non sur un marqueur formel aussi précis que l'accent de focalisation pour le français parlé.

4.1. Exemples chez l'adulte

4.1.1. En français parlé (focus étroit, suivi d'un focus large)

Si l'on s'intéresse au marquage prosodique vocal du focus étroit (FE) dans l'énoncé, seul un constituant reçoit un accent de focalisation et la syllabe correspondant à cet accent est marquée par une hauteur plus élevée, un allongement syllabique et une intensité plus forte. Dans l'exemple présenté en Figure 3, la mère demande à son fils ce qu'il a fait à l'école le jour même et celui-ci lui explique qu'il a fait des ronds et écrit son prénom. Dans l'énoncé suivant, la mère dit : « *Mais les ronds, c'était des o en fait ?* ». Dans cet énoncé, l'accent de focalisation porte sur la syllabe « o » et il est marqué au niveau acoustique par une élévation de la hauteur avec une forte augmentation de la *fo* passant de 189 à 402 Hz (+12 ½ tons) et de l'intensité relative par rapport à l'intensité moyenne de la totalité de l'énoncé (62 dB) sur le pic d'intensité à 69 dB, soit une intensité relative de 0,898), ainsi qu'un allongement de la syllabe (261 ms), qui est 60% plus longue que la précédente (157 ms, cf. Figure 3, à droite). A noter également deux autres phénomènes relevant du niveau articulatoire, la présence d'un coup de glotte juste avant la voyelle [o] (d'une durée de 87 ms) et une hyper-articulation de cette même voyelle (marquée notamment par une forte protrusion labiale très visible sur la vidéo), qui participent à la mise en saillance de la syllabe, tout comme la variation des paramètres prosodiques.

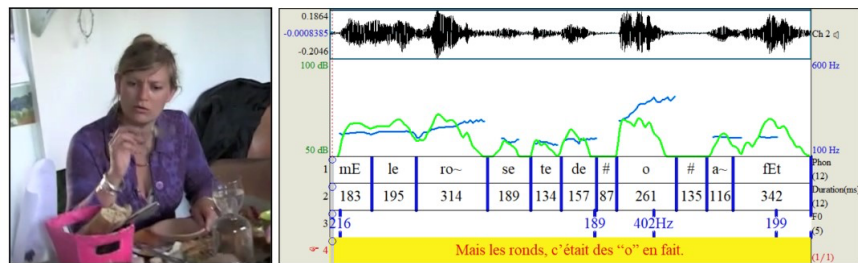


figure 3 : à gauche, la mère en train de produire « o » et à droite, représentation acoustique de l'énoncé « Mais les ronds, c'était des o en fait » (la ligne bleue représentant l'évolution de la *fo* et la ligne verte, de l'intensité).

Dans cet exemple, les variations acoustiques ne sont pas les seules à contribuer à l'effet de focalisation. Les mouvements de la tête et des mains durent plus longtemps que le focus vocal et font écho à la littérature sur la dimension gestuelle de la focalisation, selon laquelle on observe une tendance à l'alignement de l'apex du geste et de la voyelle correspondant à la cible articulatoire de la syllabe accentuée. Les mouvements des sourcils et de la tête constituent également des signaux visuels importants

pour localiser la syllabe accentuée. La mère produit en effet d'une part, un geste iconique de la main droite en traçant des cercles à travers plusieurs mouvements circulaires avec la fourchette qu'elle tient dans la même main qui produit ce geste ; d'autre part, elle associe des gestes de battement consécutifs effectués avec la tête à chaque « rond » produit avec la fourchette, et elle fronce également ses sourcils pour signaler la reformulation de l'énoncé de l'enfant dans le but de transmettre la demande de clarification. Enfin, on relève une hyper-articulation (exagération des gestes articulatoires) de la mâchoire, des lèvres et de la langue sur la syllabe accentuée, identifiable à partir du visionnage de la séquence vidéo.

Si l'on s'intéresse maintenant au marquage prosodique du FL chez l'adulte, tous les constituants majeurs de l'énoncé reçoivent une accentuation. C'est le cas dans l'exemple suivant où le père dit à son aîné qu'il va venir leur donner un cours d'acoustique. Celui-ci lui n'ayant pas compris le mot demande « *Quoi ?* » à son père qui lui explique alors de quoi il s'agit : « *L'acoustique, c'est le son* » (focus informatif) en produisant deux accents, le premier sur la syllabe finale du mot « acoustique », qui correspond au marquage habituel du thème et le second en position finale d'énoncé sur le mot « son ». Ces deux accents sont marqués différemment au niveau prosodique : le premier est marqué par une montée mélodique (de 108 à 194 Hz, soit une élévation de hauteur de 10 ½ tons, correspondant à un contour de continuation qui vient marquer la frontière du syntagme initial) et une intensité plus forte, mais la durée en revanche, n'est pas plus importante par rapport à la syllabe précédente. Le second accent est, quant à lui, marqué par une baisse de l'intonation correspondant à la modalité assertive de la phrase et un allongement de la durée de la syllabe (trois fois plus longue que la syllabe précédente), celle-ci se trouvant en position finale d'énoncé, ainsi qu'une intensité aussi forte que sur le premier accent (68 dB), avec une intensité relative de 0,926 (l'intensité moyenne étant de 63 dB).

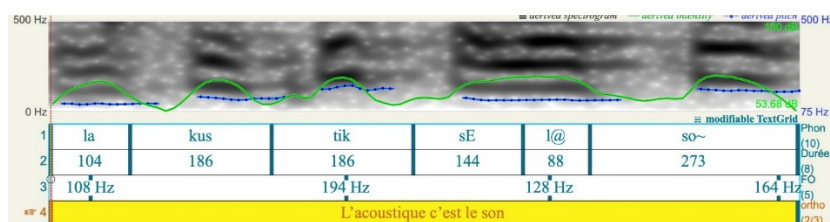


figure 4 : représentation acoustique de l'énoncé « L'acoustique c'est le son » (la ligne bleue représentant l'évolution de la *f0* et la ligne verte, l'intensité).

Au niveau gestuel, alors qu'il est en train de regarder son fils et, à mesure qu'il produit son explication, le père détourne la tête vers la gauche en regardant en direction des plats posés sur la table, puis ramène sa tête en face de son fils pour le regarder de nouveau. Le mouvement latéral gauche de la tête est synchronisé avec l'accent localisé sur la dernière syllabe d'« *acoustique* » et le mouvement latéral droit est synchronisé avec le

second accent localisé sur « son », comme on peut le visualiser sur les trois images ci-dessous (cf. Figure 5).



figure 5 : les trois phases du mouvement de tête du père (fixation du fils, mouvement latéral vers la gauche, puis vers la droite et retour sur la position d'origine).

4.1.2. En LSF

Il nous semble intéressant de mettre en regard l'extrait précédent (Figure 5) en français parlé et l'extrait suivant (Figures 6 et 7) en LSF. La mère introduit un nouveau sujet dans la discussion avec sa fille aînée, et doit en particulier attirer son attention dans la mesure où la disposition autour de la table favorise les binômes participatifs [Mère-Enfant la plus jeune] et [Père-Enfant aînée]. C'est pourquoi l'énoncé de la mère commence par ce qui pourrait être traduit en français par 'Oh, Charlotte, j'avais oublié, j'ai une question pour toi'. Puis, elle introduit un FL, englobant l'ensemble de la question, avec deux points d'attention (quand et où), que nous avons retenus dans les items focalisés en (2).

(2) LSF NOM-SIGNE(Ea) OUBLIE QUESTION **DIMANCHE!**,
SAIT VA **NOM-SIGNE(asso)**!
trad. 'Oh, Charlotte, j'avais oublié, j'ai une question pour toi : dimanche
(prochain), tu sais qu'on va à (l'association) Nom-Signe ?'

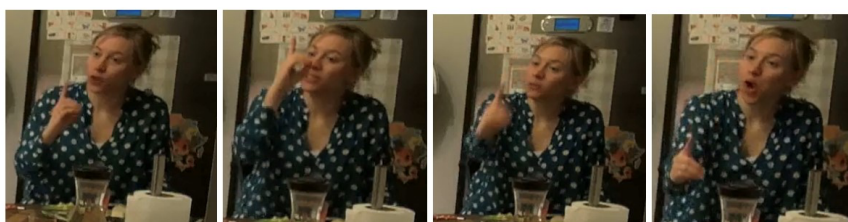


figure 6 : [DIMANCHE-PROCHAIN!]



figure 7 : [ASSOCIATION-DLS!]

L'accent de focalisation sur le signe [DIMANCHE-PROCHAIN] (Figure 6) est marqué par une amplitude relativement forte du mouvement du bras, suivi d'une légère tenue, ainsi que d'une labialisation hyper-articulée, et l'accent de focalisation sur le signe [NOM-ASSOCIATION] (Figure 7) est marqué par une amplitude relativement forte du mouvement de la main et avant-bras, souligné par une hyper-extension des doigts, suivi d'une tenue et de sa position finale dans l'énoncé.

4.2 Exemples chez l'enfant

4.2.1. En français parlé (focus large, suivi d'un focus étroit)

Le marquage de la focalisation annoté ici chez l'enfant est le plus souvent étroit (une seule syllabe focalisée à l'aide des trois paramètres prosodiques : hauteur, durée, intensité relative) même si quelques-uns sont larges (plusieurs syllabes accentuées, toujours dans le cas de la négation, avec quelques fois l'apparition d'un pattern de scansion très spécifique et très rythmique, que nous décrirons plus loin dans la partie 4.3). Dans l'exemple suivant, le père donne à son enfant (3 ans) des morceaux de melon, que celui-ci vient de réclamer. Après un petit jeu entre le père et l'enfant pour que le père le serve beaucoup plus, l'enfant dit « *Non c'est pas beaucoup* » (FL) puis répète « *C'est pas beaucoup* » (FE). L'accent de focalisation produit sur l'adverbe "pas" est beaucoup plus saillant dans le second énoncé (cf. Figure 8). En effet, dans le premier énoncé, on relève une montée mélodique de 2 ½ tons (en passant de 322 à 368 Hz), alors que dans le second, elle est de 5 ½ tons (en passant de 324 à 433 Hz). On note le même phénomène pour les deux autres paramètres prosodiques que sont la durée (passant de 328 à 372 ms) et l'intensité (passant de 76 à 81 dB, soit une légère montée en intensité relative de 0,894 à 0,9). On observe donc une gradation prosodique dans le second énoncé concernant le marquage de l'accent de focalisation, accompagnant la restriction du focus sur l'adverbe « pas ».

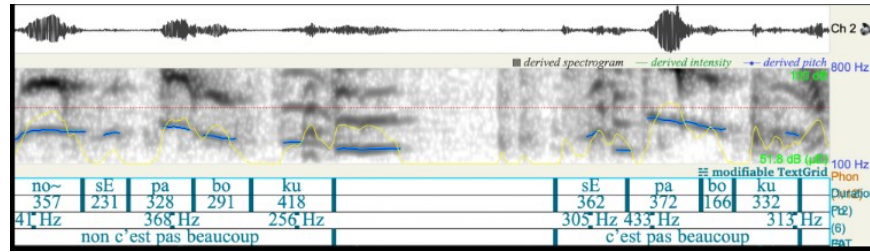


figure 8 : représentation acoustique des deux énoncés « *Non c'est pas beaucoup* » et « *C'est pas beaucoup* » (la ligne bleue représentant l'évolution de la F0 et la ligne jaune, de l'intensité).



figure 9 : à gauche, montée de la main gauche avec le poing fermé à hauteur de la bouche de l'enfant lorsque celui-ci produit le "non" dans le premier énoncé.



figure 10 : de nouveau, montée plus marquée de la main gauche à hauteur des yeux sur le "pas" du second énoncé.

Sur le plan corporel, dans les deux énoncés, il est intéressant de noter le mouvement du bras gauche de l'enfant (cf. Figure 9). Dans le premier énoncé, le bras gauche avec le poing fermé s'élève pendant la production du mot « *non* » jusqu'à couvrir sa bouche, suivi d'une légère élévation de la main pendant la production du mot *pas* (en synchronisation avec la parole). Dans le second énoncé (cf. Figure 10), la main qui était restée en face de la bouche s'étend vers l'avant sur le mot « *c'est* » puis se rétracte pour se placer au niveau de l'œil gauche de l'enfant sur le mot « *pas* », pour redescendre au niveau de la bouche à la fin de l'énoncé, toujours en synchronisation avec la parole.

4.2.2. En LSF

On observe quelques extraits où l'enfant produit un énoncé avec plusieurs accents de focalisation reliés par un contour prosodique. Dans l'exemple (3), la fille aînée demande à sa mère de lui servir à boire. Elle

hésite et se reprend plusieurs fois en jouant avec la forme des signes de fruit, pour dire finalement qu'elle souhaite du jus de pomme.

(3)

LSF STP JUS PECHE/POMME, geste BÊTE, STP JUS
ORANGE/POMME, geste COGNE **STP! JUS! POMME!** geste
Trad. 'S'il te plaît, du jus de pem..., euh non, je suis bête, s'il te plaît du jus de
por... Ah ! je suis vraiment bête, s'il te plaît, du jus de pomme !!!'



figure 11 : [S-IL-TE-PLAIT]



figure 12 : [JUS]



figure 13 : [POMME]

Les marqueurs de focalisation sur [S-IL-TE-PLAIT] sont un large hochement de la tête accompagnant un mouvement manuel relativement ample sur le signe [JUS], on observe un deuxième hochement de tête, souligné par le basculement de la tête en arrière au début du signe et un mouvement manuel relativement ample. Sur le signe [POMME], on observe un écho et crescendo sur la séquence avec un troisième hochement de tête relativement ample, associé à une extension du buste, une hyper-articulation manuelle et labiale et la vocalisation [pɔm]. L'effet de crescendo provient d'une augmentation progressive de l'intensité des marqueurs de la focalisation sur les trois occurrences. L'enfant est en effet agacé de ne pas réussir à s'exprimer et s'emporte sur la fin de sa phrase (ou joue à l'être).

En français parlé comme en LSF, nous avons donc trouvé et décrit des catégories de marqueurs (Figure 14) rapportées dans la littérature sur l'expression prosodique de la focalisation, qu'elle concerne le FE ou le FL. Mais notre attention a également été attirée par des séquences (parmi les plus longues) qui forment un schéma prosodique spécifique, dans les deux langues, dans les productions des adultes et des enfants. Il s'agit de séries de temps forts, coïncidant avec des syllabes vocales *ou gestuelles* accentuées, qui forment ce que nous appelons un « patron de scansion ».

- | | |
|--|--|
| <ul style="list-style-type: none"> • Français <ul style="list-style-type: none"> • Prosodie vocale <ul style="list-style-type: none"> • Proéminence mélodique • Pic d'intensité • +/- allongement syllabique • Gestes <ul style="list-style-type: none"> • Mouvements de la tête (en avant ou inclinaison) • Élévation ou abaissement des sourcils • pointage • Mouvement des mains en avant | <ul style="list-style-type: none"> • LSF <ul style="list-style-type: none"> • Prosodie gestuelle <ul style="list-style-type: none"> • contrastes d'amplitude (mains, tête, face, mouvements du corps) • contraste dans la localisation (proéminence dans l'espace de signation) • accélération • tenue • bouche en « hyper-articulation » |
|--|--|

figure 14 : Inventaire des marqueurs de focalisation

4.3. Patron de scansion

4.3.1. Exemple chez l'adulte

En français parlé

Dans l'extrait suivant (Figure 15), la mère récupère les cure-dents utilisés par les enfants assis à côté d'elle pour manger du melon. Si l'enfant à sa gauche donne tout de suite son cure-dent à la mère, l'enfant à sa droite ne semble pas vouloir y donner suite. La mère répète alors sa demande avec une intensité plus accentuée et scande chaque mot de son énoncé tout en marquant notamment l'attaque de chaque unité linguistique.

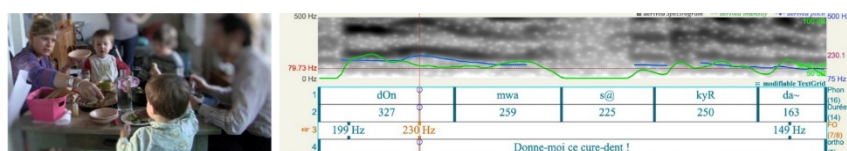


figure 15 : à gauche, marqueurs non verbaux accompagnant le focus large produit par la mère et à droite, représentation acoustique de l'énoncé « Donne-moi ce cure-dent » (la ligne bleue représentant l'évolution de la *fo* et la ligne jaune, de l'intensité).

Le pattern de scansion porte sur les trois premières syllabes : « *Donne-moi-ce* » dont la durée syllabique, comme dans l'exemple précédent, est allongée (376-296-271 ms), mais cette fois-ci, avec un effet inverse de réduction de la durée au cours de la progression, cette réduction étant encore plus marquée sur les deux dernières syllabes (188-177 ms). Sur le plan gestuel, on relève l'implication de tous les articulateurs (cf. Figure 13). Premièrement, la mère tourne son buste vers l'enfant, et son regard est dirigé vers ce dernier pour établir un contact visuel. On remarque également un léger froncement des sourcils qui exprime son impatience. En même temps, si sa main gauche tient le bol où sont déposés les cure-dents récupérés précédemment, sa main droite est configurée de sorte à produire un pointage associé au démonstratif « *ce* » qui anticipe le syntagme plein « *cure-dent* ». En outre, on relève des gestes de battement du haut vers le bas sur l'axe vertical, produits avec la tête en correspondance avec l'attaque initiale de chaque mot. Ainsi, plus le nombre de constituants comportant une accentuation est important (ici, le phénomène de scansion porte sur plusieurs syllabes successives), plus les productions gestuelles sont marquées, rendant compte du fait que ce comportement est pleinement multimodal.

En LSF

Le phénomène de scansion en LSF a été étudié dans le registre poétique enfantin (Blondel, 2000) ou dans l'expression de slogans (Blondel & Schmitt, à paraître), où sont mis en lumière des patrons rythmiques alternant temps forts et temps faibles. Les temps forts correspondent aux syllabes accentuées 'par défaut', autrement dit la partie des mouvements des signes qui portent leur dimension significative, associées éventuellement à une forme d'accent d'emphase ; les temps faibles correspondent par défaut aux mouvements non accentués de transition intra- ou inter-signe, ou aux signes qui sont moins saillants pour différentes raisons (phonologique, sémantique, ou en raison de leur place dans la structure informationnelle).

L'exemple (4) est extrait d'une discussion entre la mère et sa plus jeune fille. L'enfant verse l'eau qui vient de lui être servie dans le verre de sa mère, car elle veut du sirop de menthe. La mère est surprise (« *Hé, mais pourquoi tu fais ça ? Qu'est-ce que tu fais ?* »), s'appuie sur l'avis paternel et rappelle à sa fille qu'elle n'a pas le droit d'avoir du sirop à table.

(4)
 Accent – –’ – –’
 LSF geste (PA)PA! DIT geste-négation!
 Trad. ‘Hé, Papa a dit non !

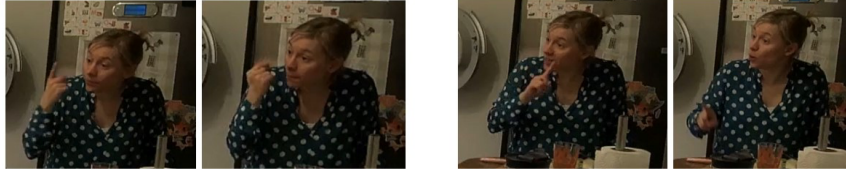


figure 16 : [PAPA!] et [geste-négation!]

L'énoncé se caractérise par ce patron de scansion que nous avons décrit précédemment en français parlé. Il se compose d'un mouvement de transition vers la position initiale du signe [PAPA], suivi d'une tenue en début de signe, puis de la syllabe accentuée par lehaussement d'épaule et des sourcils, associé à un mouvement de la tête vers l'avant ; précédé du signe [DIT] relativement escamoté, le geste ou signe [geste-négation] est marqué par une amplitude relativement large et une hyper-articulation labiale. L'énoncé dessine donc le patron prosodique [() _ _ _] en alternant temps faibles et temps forts.

4.2.2 Exemples chez l'enfant

En français parlé

Chez l'enfant, les focus produits sont tous accompagnés de configurations gestuelles, ce qui montre à quel point la multimodalité est un support chez l'enfant, dont le développement langagier est toujours en évolution, pour communiquer et pour signaler certains segments de son discours (Morgenstern, 2014 ; ce numéro).

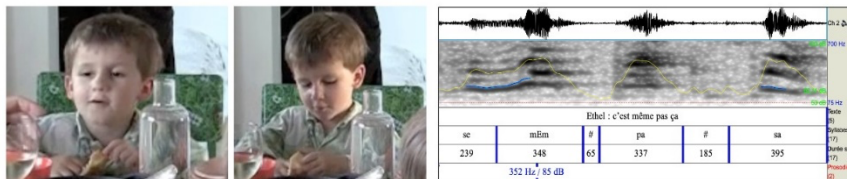


figure 17 : à gauche, marqueurs non verbaux accompagnant le focus large produit par l'enfant et à droite, représentation acoustique de l'énoncé « C'est même pas ça ! » (la ligne bleue représentant l'évolution de la *f₀* et la ligne jaune, de l'intensité).

L'exemple de la Figure 17 illustre un FL dans l'énoncé précédent « *Non, c'est même pas ça* » où l'enfant refuse la correction apportée par son grand

frère, et un FE lorsque l'enfant répète une seconde fois les mêmes mots qu'il vient de produire pour renforcer son désaccord : « *C'est même pas ça !* » avec une montée mélodique sur l'adverbe « *même* » (de 256 à 346 Hz, soit une montée de 6 ½ tons, plus marquée que la première montée mélodique réalisée au même endroit dans le premier énoncé avec une montée de 5 ½ tons, de 286 à 369 Hz). C'est dans cet énoncé qu'émerge un pattern de scansion (cf. représentation acoustique, Figure 15, à droite), avec des durées syllabiques plus importantes que la normale (239-348 # 337 # 395 ms), mais à la différence des deux adultes, Azul – le prénom du cadet de la fratrie - introduit deux pauses (# : la première d'une durée de 65 ms et la seconde, de 185 ms). Il y a donc là aussi un effet de gradation avec une tendance à l'allongement syllabique plus on progresse dans l'énoncé et la présence de ces deux pauses contribue à individualiser davantage encore chacune de ces syllabes et à renforcer l'effet rythmique de la scansion. Il est également possible que l'enfant ait encore besoin de produire des pauses pour renforcer l'effet de scansion, alors que les adultes n'en ont plus besoin car ils ont un meilleur contrôle phonatoire.

Ces deux focus ayant une fonction contrastive, ils ne sont pas produits de la même façon du point de vue multimodal. Dans le cas du FE sur l'adverbe « même » dans le premier énoncé, l'enfant, qui est en train de manipuler le bout de pain dans ses mains tout en le regardant, soulève d'abord la tête pour diriger son regard vers son grand frère et se penche simultanément vers l'avant, engageant ainsi la totalité de son corps pour réagir à la correction de celui-ci. Lorsque l'enfant répète son énoncé, un FL se met en place au moment de la scansion de chaque mot (c'est *û* même # pas # ça), mettant en relief le contraste créé avec ce que son interlocuteur dit et à partir de trois groupes intonatifs. Sur le plan gestuel, on peut voir que l'enfant recule légèrement et baisse sa tête pour fixer son regard sur le bout de pain qu'il tient toujours dans ses mains (combinaison « Regard-Tête-Corps »), ce qui montre comment l'enfant exprime son opposition, où dimensions prosodique et gestuelle se combinent pour véhiculer le message de contraste par rapport à ce que son grand frère lui a dit.

En LSF

L'exemple (5) correspond à l'échange à propos du sirop de menthe vu du côté de l'enfant la plus jeune qui, après avoir versé l'eau que sa mère lui avait servi dans le verre de celle-ci, réclame du sirop de menthe, puis proteste vivement, suite au refus de sa mère (présenté en 4).

(5a)
 Accent (--) _' _ _'
 LSF **VEUX!** **MENTHE!**
 Trad. ' [a] veux d' la menthe !



figure 18 : [VEUX] et [MENTHE]

(5b)
 Accent (-) - ' - - ' -
 LSF (trans) PTx! EAU[esc] geste-négation!
 Trad. 'alors ça, l'eau, j'en veux pas !'."



figure 19 : [PT->eau]

et

[geste-négation]

Les deux énoncés de l'enfant, séparés par une pause pendant laquelle elle écoute-regarde sa mère, se caractérisent par un patron rythmique similaire de scansion. Les mouvements manuels de l'enfant des signes accentués sont perçus comme accélérés, d'une amplitude relativement plus large que celle des signes en contexte, accompagnés de mouvements de tension du buste et des épaules, suivis d'un relâchement brusque. On observe également une hyper-articulation des articulateurs de la face et l'oralisation des signes accentués. Notons par ailleurs, que dans cet exemple du « sirop de menthe », le patron de scansion se construit aussi dans l'interaction enfant-mère-enfant, sur un motif rythmique binaire, avec le signe ou geste de négation qui représente le pic prosodique en écho chez la mère et la fille.

5. Discussion-Conclusion

En LSF, nous retrouvons dans les exemples présentés ce que les différents auteurs cités en état de l'art ont relevé pour la focalisation prosodique et ses manifestations à travers les articulateurs manuels et non manuels. Nous observons également, comme Crasborn et van der Kooij (2013) ou Navarrete-González (2021), que des phénomènes non manuels peuvent être distribués sur différents articulateurs selon les contraintes de

la distribution de l'item focalisé. Autrement dit, ce n'est pas un haussement de sourcil ou une avancée du buste en tant que tels qui marquent le focus, mais davantage le patron que le phénomène créé, par contraste avec ce qui précède et ce qui suit, rendant saillant ainsi l'item sous focus. Les marqueurs de prosodie gestuelle en LSF relevés ici sont le résultat i) de contrastes d'amplitude des mouvements des mains, de la tête, des articulateurs de la face, et du buste, ii) de contrastes dans la localisation des signes (les rendant proéminents dans l'espace de signation), iii) de phénomènes rythmiques avec accélération ou tenue, iv) de labialisations hyper-articulées.

En français parlé, un certain nombre de marqueurs liés à la focalisation ont été mis au jour à la fois au niveau vocal et au niveau gestuel. Au niveau vocal, la focalisation est le plus souvent associée à une variation des paramètres prosodiques affectant une ou plusieurs syllabes, c'est-à-dire la hauteur, les allongements syllabiques la plupart du temps et l'intensité, associés aux paramètres acoustiques de *f₀*, de durée et d'amplitude. Zenghui et al. (2020) font le même constat dans leur travail pour identifier quelles ressources sont utilisées par les locuteurs pour produire des séquences focalisées ou des questions déclaratives. En l'occurrence, ils relèvent que le marquage du focus s'appuie sur des modulations concomitantes de l'étendue de la hauteur et de la durée des segments syllabiques dans différents types d'énoncés. Au niveau gestuel (manuel et non manuel), toujours en français parlé, lorsque l'on rencontre un focus large ou étroit, on observe que la focalisation est caractérisée par des mouvements de plusieurs articulateurs qui participent à la scansion et à l'accentuation de certains segments de l'énoncé et ce, variant parfois sur le plan qualitatif selon le profil (adulte/enfant) du locuteur. Ainsi, si le focus est mis en place par l'intermédiaire de plusieurs articulateurs mobilisés conjointement au marquage vocal, et moins par la mobilisation d'un seul articulateur, pris individuellement, on peut certainement conclure que la focalisation en français parlé, qu'elle soit étroite ou large, est un phénomène multimodal qu'il est nécessaire d'observer dans un « fait » plus large, lui aussi multimodal, à savoir le langage en interaction.

Sur la base de cet inventaire de marqueurs concernant les focus étroits ou larges, et de notre intérêt pour des patrons prosodiques plus larges, tels que les schémas prosodiques avec scansion, nous observons des correspondances intéressantes entre les langues parlées (y compris leur dimension gestuelle) et les langues des signes. Nous avons ainsi constaté une amplification des valeurs des paramètres qui peut également être exprimée par une hyper-articulation (vocale ou plus largement gestuelle). Ce phénomène d'hyper-articulation a été décrit pour les langues parlées, y compris en tenant compte de la dimension visuelle de la prosodie vocale : Ambrazaitis et House (2022) y réfèrent en partie lorsqu'ils proposent la *Cumulative Cue Hypothesis*, selon laquelle il y aurait une relation entre la proéminence prosodique du mot - exprimée par la montée de l'accent - et une accumulation de marqueurs gestuels impliquant notamment la production simultanée des mouvements de battement de la tête et des

mouvements des sourcils. Mais cette hyper-articulation fait écho également à l'observation de Crasborn et van der Kooij (2013 : 516) pour l'expression d'un focus contrastif en LS, qui se manifeste par un 'effort articulatoire accru' ("*appears to lead to increased articulatory effort*"). Nous avons également remarqué que, pour qu'un item soit perçu comme focalisé, il doit s'inscrire systématiquement dans une distribution contrastive, autrement dit, une syllabe est perçue comme accentuée parce qu'elle est entourée de tenues, allongements, intensité plus faible, etc. L'observation des patrons qui se dessinent à partir de séries d'événements (répartis dans un focus large) ne contredit pas l'importance de la valeur relative et contrastive du marquage de la focalisation : la régularité d'un patron apparaît par contraste avec l'irrégularité de son contexte. Par ailleurs, le fait que l'item soit perçu comme plus ou moins marqué semble être lié non seulement à la valeur d'un paramètre en tant que telle, mais aussi à l'alignement et à la synchronisation entre les différents marqueurs, qui peuvent contribuer à un renforcement de l'accentuation.

Appréhender les phénomènes prosodiques en circulant entre les données 'langues parlées' et 'langues des signes' nous permet de réfuter, pour les LS, l'approche dichotomique (évoquée en 2.2.1) : intonation non manuelle vs. rythmique manuelle. En effet, nous prenons en compte l'aspect dynamique des flux (sonore ou gestuel) et considérons que les paramètres de durée, intensité/amplitude, fréquence sont compatibles avec une circulation entre les langues, entre les modalités et s'appliquent aux mouvements des différents articulateurs, quelle que soit leur taille. En d'autres termes, non seulement nous nous intéressons aux liens entre la prosodie vocale (*speech prosody*) et la gestualité du corps dans son ensemble (*body gesture*) (Esteve-Gibert et Guellai, 2018 : 1, entre autres), mais nous proposons d'aller plus loin en incluant, dans l'étude de la prosodie des langues parlées, ses deux modalités, sonore et gestuelle. Cette approche originale nous permet de prendre en compte la dynamique de la gestualité associée à la parole vocale, en nous intéressant, par exemple, à la présence d'un pointé manuel à la fois pour sa fonction déictique globale et pour son potentiel à accueillir ou non un phénomène accentuel dans la structure de son mouvement. Nous abordons le mouvement dans son déroulement continu et non comme une succession de positions (même si la représentation d'un instantané permet de rendre visible le déplacement, comme le montrent les illustrations dans cet article).

A l'instar de travaux consacrés à la synchronisation entre parole vocale et gestualité associée (Pouw et al., 2020, par exemple), nous explorons actuellement les possibilités de mesure des phénomènes fins de (dé)synchronisation entre mouvements des articulateurs (qu'ils soient impliqués la production vocale, manuelle ou plus généralement corporelle). La dimension écologique et interactionnelle des données observées ici nous empêche de recourir à des dispositifs intrusifs et nous encourage à exploiter les outils de mesure appuyés sur la lecture des enregistrements vidéos (Catteau et Blondel, 2023 ; Catteau et Bianchini, ce volume). Là encore, le caractère innovant de notre démarche est

d'appliquer ce protocole de mesure des mouvements dans les deux types de données : LSF et français parlé (appréhendé dans sa multimodalité).

L'ensemble des observations effectuées dans le panel présenté dans cet article demande à être confronté à un corpus plus étendu et plus varié, afin notamment de pouvoir observer ce qui relève de la perspective développementale, et comparer les patrons observés chez les plus jeunes des enfants du corpus DinLang à ceux observés chez les plus âgés d'entre eux. En effet, dans un état de l'art relativement récent, Esteve-Gibert et Guellaï (2018) soulignent l'importance, chez les tout-petits (entendants), de la perception de l'alignement temporel et de la congruence entre gestualité et vocalité : selon les travaux recensés, cette intégration audiovisuelle contribue à la segmentation et au traitement de l'input, à l'organisation de l'information et à l'attribution d'intentions communicationnelles spécifiques chez l'interlocuteur par le tout-petit. De plus, les enfants semblent développer progressivement des patrons prosodiques audio-gestuels 'cibles', se rapprochant de ceux des adultes. De plus amples explorations des productions sont donc nécessaires pour confirmer la chronologie de ce processus d'acquisition. Dans une étude longitudinale appuyée sur des données bilingues LSF-français et LSF (Limousin et Blondel, 2010), nous avons formulé une hypothèse analogue : avec la maîtrise progressive de la langue, les petites signeuses s'approprient peu à peu les patrons prosodiques présents dans les productions parentales et modulent avec plus de précision les contours prosodiques dans leurs productions. Par ailleurs, nos analyses sur le marquage gestuel de la focalisation en français parlé ont montré ici que les adultes ont tendance à créer un focus multimodal par l'utilisation d'un plus grand nombre d'articulateurs, en comparaison avec l'enfant. Cela montrerait probablement que la correspondance entre la composante prosodique vocale et sa contrepartie gestuelle est toujours en cours de stabilisation chez le jeune locuteur (Esteve-Gibert et Guellaï, 2018 ; Hübscher et Prieto, 2019). Dans les données étudiées ici, nous avons illustré à travers les échanges entre des adultes et des enfants relativement jeunes (entre 2 et 9 ans), la production de patrons prosodiques analogues, comme celui de la scansion, se faisant même parfois écho dans la co-construction de l'interaction (exemples 4 et 5).

Toutes ces observations nous invitent à réfléchir sur la manière dont la gestualité (manuelle et non manuelle) est mobilisée par les locuteurs entendants et à reconsidérer les langues des signes et les langues parlées, avant tout, comme des langues incarnées. Ainsi nous appelons de nos vœux une typologie linguistique, y compris au niveau prosodique, qui prendrait en compte l'inter-modalité.

Références bibliographiques

Ambrazaitis, G., & House, D., 2022, Probing effects of lexical prosody on speech-gesture integration in prominence production by Swedish news presenters.

- Laboratory Phonology*, 13(1:15), p. 1-35. <https://doi.org/10.16995/labphon.6430>
- Astésano C., Magne C., Morel M., Coquillon A., Espesser R., Besson M. & Lacheret-Dujour A., 2004, Marquage acoustique du focus contrastif non codé syntaxiquement en français, *Actes des Journées d'Étude sur la Parole (JEP)*, Fès, Maroc, p. 1-4.
- Avanzi M. & Martin P., 2007, L'intonème conclusif : une fin (de phrase) en soi ? *Cahiers de linguistique française* 28, p. 247-258.
- Bagou O., 2001, Validation perceptive et réalisations acoustiques de l'implication emphatique dans la narration orale spontanée, *Cahiers de Linguistique Française* 23, p. 39-59.
- Beysade C., Hemforth B., Marandin J.-M. & Portes, C., 2009. Prosodic Markings of Information Focus in French, *Interface Discours & Prosodie*, Paris, 9–11 Septembre.
- Boersma P. & Weenink D., 2023, Praat: doing phonetics by computer [Computer program]. Version 6.3.09, retrieved 2 March 2023 from <http://www.praat.org/>
- Boyes-Braem P., 1999, Rhythmic temporal patterns in the signing of deaf early and late learners of German Swiss Sign Language, *Language and Speech* 42 2&3, p. 177-208.
- Blondel M., 2000, *Poésie enfantine dans les langues des signes : Modalité visuo-gestuelle versus modalité audio-orale*, Thèse de doctorat. Université François Rabelais, Tours.
- Blondel M., 2003, Saillance linguistique dans une langue à modalité gestuelle, in J. François & A. Lacheret (eds), *Fonctions et moyens d'expression de la focalisation à travers les langues.*, Peeters, pp. 187-203., Mémoires de la Société de Linguistique de Paris, Tome XIII.
- Blondel M., 2020, *Les langues des signes, des langues incarnées. Linguistique, Habilitation à Diriger des Recherches*. Université Paris 8.
- Blondel M. & Le Gac D., 2007, Entre parenthèses... y a-t-il une intonation en LSF ? *Silexicales*, 5, p. 1-16.
- Blondel M. & Schmitt P., 2022, Signer à l'unisson : ce que révèlent les « slogans pour l'œil », *Cahiers de littérature orale* [En ligne], 91-92 | 2022, mis en ligne le 28 février 2024, consulté le 04 mai 2024. URL : <http://journals.openedition.org/clo/11575>
- Brown L. & Prieto P., 2021. Gesture and Prosody in Multimodal Communication, in M. Haugh, D. Z. Kádár & M. Terkourafi (eds), *The Cambridge Handbook of Sociopragmatics*, Cambridge University Press, p. 450-453.
- Cappeau P. & Hanote S., 2012, *Focalisation(s). Saillance dans les langues : lexicale, syntaxe, prosodie*, Presses Universitaires de Rennes.
- Catteau F., 2020, *Traduire la poésie en langue des signes : l'empreinte prosodique lors du changement de modalité*, Thèse de doctorat. Université Paris 8.
- Catteau F. & Blondel M., 2023, Pause and hold in LSF: how to control one's gestural fluency? Communication affichée, *GESPIN*, Sep 2023, Nimegue, Netherlands.
- Catteau F. & Bianchini C. (ce volume), Mise en place d'un protocole d'identification de traits articulatoires communs à la gestualité co-verbale du français et à la Langue des Signes Française (LSF) : application au geste épistémique.

- Crasborn O. & van der Kooij E., 2013, The phonology of focus in Sign Language of the Netherlands, *Journal of Linguistics* 49(3), p. 515-565. <http://www.jstor.org/stable/24583284>
- Delattre P., 1966, *Studies in French and Comparative Phonetics*, The Hague: Mouton.
- De Paolis B.M., Santiago F. & Andorno C.M., 2022, Syntaxe ou prosodie ? Une étude préliminaire sur l'expression de la focalisation étroite par les apprenants italo-phones de français L2, *34e Journées d'Études sur la Parole*, Île de Noirmoutier.
- Di Cristo A., 1999a, Le cadre accentuel du français contemporain : essai de modélisation (première partie), *Langues* 2(3), p. 184-205.
- Di Cristo A., 1999b, Le cadre accentuel du français contemporain : essai de modélisation (seconde partie). *Langues* 2(4), p. 258-267.
- Di Cristo A., 2000, Vers une modélisation de l'accentuation du français (seconde partie), *Journal of French Language Studies* 10/1, p. 27-44.
- Dohen M., Løevenbruck H., Roustan B. & Vilain C., 2011, Pointage manuel et vocal : corrélats sensorimoteurs et neurocognitifs, coordination geste manuel / parole, *Faits de Langues* 37, p. 97-116.
- Dohen M. & Løevenbruck H., 2009, Interaction of audition and vision for the perception of prosodic contrastive focus, *Language & Speech* 52(2-3), p. 177-206.
- Dohen M., & Løevenbruck H., 2005, Audiovisual Production and Perception of Contrastive Focus in French: a multispeaker study, *Interspeech/Eurospeech*, Lisbonne, Portugal, p. 2413-2416.
- Dohen M., Løevenbruck H. & Harold H., 2006, Visual Correlates of Prosodic Contrastive Focus in French: Description and Inter-Speaker Variability. *Speech Prosody*, May, Dresden, Germany. p. 221-224.
- Esteve-Gibert N. & Guellaï B., 2018, Prosody in the auditory and visual domains: A developmental perspective, *Frontiers of Psychology*, 9:338. doi: 10.3389/fpsyg.2018.00338
- Esteve-Gibert N., Løevenbruck H., Dohen M. & D'Imperio M., 2022, Pre-schoolers use head gestures rather than prosodic cues to highlight important information in speech, *Developmental Science*, 25(1). <https://doi.org/10.1111/desc.13154>
- Fenlon J., Cooperrider K., Keane J., Brentari D. & Goldin-Meadow S., 2019, Comparing sign language and gesture: Insights from pointing, *Glossa: A Journal of General Linguistics*, 4, p. 1-26.
- Ferré G., 2019, *Analyse de Discours Multimodale. Gestualité et prosodie en discours*, Grenoble, UGA Éditions.
- Ferré G., 2018, Gesture/speech integration in the perception of prosodic emphasis, *Actes du colloque international Speech Prosody*, Poznan, Pologne, p. 35-39.
- Ferré G., 2003, Discursive, prosodic and gestural analysis of focalisation pauses in british english. *Interfaces prosodiques Proceedings*, p. 265-270. Université de Nantes.
- Féry C., 2001, Focus and phrasing in French, in C. Féry & W. Sternefeld (eds.), *Audiatur Vox Sapientiae. A Festschrift for Arnim von Stechow (Studia grammatica*, vol. 52, Berlin: Akademie Verlag, p. 153-181.

- Godement-Berline R., 2018, *La focalisation prosodique dans la parole interprétée en français*, Thèse de Doctorat. Université Sorbonne Paris Cité.
- Graf H. E., Cosatto E., Strom V. & Huang F. J., 2002, Visual prosody: facial movements accompanying speech, *Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition*, Washington, DC, USA, p. 396-401, doi: 10.1109/AFGR.2002.1004186.
- Granström B. & House D., 2005, Audiovisual representation of prosody in expressive speech communication, *Speech Communication*, 46(3-4), p. 473-484.
- Granström B., House D. & Lundeberg M., 1999a, Eyebrow movement as a cue to prominence. *Psychology*.
- Granström B., House D. & Lundeberg M., 1999b, *Prosodic Cues in Multimodal Speech Perception*, ICPHS99, San Francisco.
- Groussier M.L. & Rivière C., 1996, *Les mots de la linguistique. Lexique de linguistique énonciative*, Paris : Ophrys.
- Grosjean F. & Lane, H., 1977, Pauses and syntax in American Sign Language, *Cognition* 5, p. 101-117.
- House D., Beskow J. & Grandström B., 2001, Timing and interaction of visual cues for prominence in audiovisual speech perception, *Proceedings of Interspeech*, september 3-7. DOI: 10.21437/Eurospeech.2001-61
- Hübscher I. & Prieto P., 2019, Gestural and prosodic development act as sister systems and jointly pave the way for children's sociopragmatic development, *Frontiers in Psychology* 10, Article 1259. <https://doi.org/10.3389/fpsyg.2019.01259>
- Jun S.-A. & Fougeron C., 2000, A phonological model of French intonation, in A. Botinis (eds), *Intonation: analysis, modeling and technology*, Dordrecht: Kluwer Academic Publishers, p. 209-242.
- Katz J. & Selkirk E., 2011, Contrastive focus vs. discourse-new: Evidence from phonetic prominence in English, *Language* 87(4), p. 771-816.
- Kendon A., 1972, Some Relationships between Body Motion and Speech, in A. W. Siegman & B. Pope (eds), *Studies in Dyadic Communication*, Elmsford, NY: Pergamon Press, p. 177-210.
- Kooij E., van der, Crasborn O. & Emmerik W., 2006, Explaining prosodic body leans in Sign Language of the Netherlands: pragmatics required, *Journal of Pragmatics* 38, p. 1598-1614.
- Krahmer E. J., Ruttkay Z., Swerts M. G. J. & Wesselink J. W., 2002, Pitch, eyebrows and the perception of focus, *Proceedings of the Speech Prosody 2002 Conference, Aix-en-Provence, France, April, 11-13*, p. 443-446.
- Lacheret A., 2003, Focalisation et circonstance : que nous dit la prosodie du français parlé ? Mémoires de la Société de Linguistique de Paris, *Fonctions et moyens d'expression de la focalisation à travers les langues*. Nouvelle série tome 13, p. 137-160.
- Lacheret-Dujour A. & François J., 2003, Fonctions et moyens d'expression de la focalisation à travers les langues [avant-propos]. *Mémoires de la Société de Linguistique de Paris, Fonction et moyens d'expression de la focalisation à travers les langues*. Nouvelle série tome 13, p. 7-12.
- Lambrecht K., 1994, *Information structure and sentence form: Topic, focus and the mental representations of discourse referents*. Cambridge University Press.

- Limousin F. & Blondel M., 2010, Prosodie et acquisition de la langue des signes française : Acquisition monolingue LSF et bilingue LSF-français, *LIA* 1:1, p. 82-109.
- Linell P., 2009, *Rethinking Language, Mind and World dialogically: Interactional And Contextual Theories Of Human Sense-Making*. Charlotte, NC: Information Age Publishing.
- Loehr D., 2014, Gesture and prosody. In C. Müller, A. Cienky, E. Fricke, S. Ladewig, D. McNeill & J. Bressems (eds.), *Body - Language - Communication* 2, 1381-1391. München/Boston, De Gruyter Mouton, 1381-1391.
- Lombart C., 2021a, Exploring the interaction between co-speech gestures and speech prosody in the marking of contrastive focus in French: a pilot study, *4th Phonetics and Phonology in Europe* (PaPE2021).
- Lombart C., 2021b, Au croisement des ressources orales, gestuelles et signées : Comparaison de la prosodie du français et de la LSF, *Papers of the LSF* (Vol 15).
- Mondada L., 2017, Nouveaux défis pour l'analyse conversationnelle : L'organisation située et systématique de l'interaction sociale, *Langage et société* 160-161(2-3), p. 181-197.
- Moody B., 1983, *La langue des signes, Tome 1. Introduction à l'histoire et à la grammaire de la langue des signes*. Paris : IVT.
- Morel M.-A., Danon-Boileau L., 1998, *Grammaire de l'intonation : L'exemple du français*, Paris, Ophrys.
- Morgenstern A., 2014, Children's Multimodal Language Development, in C. Fäcke (eds), *Manual of Language Acquisition*. De Gruyter, p. 123-142.
- Morgenstern A. & Caët S., 2021, *Signes en famille* [Corpus]. ORTOLANG (Open Resources and TOols for LANGuage) - www.ortolang.fr, v1.1, <https://hdl.handle.net/11403/signes-en-famille/v1.1>.
- Morgenstern A., Caët S., Debras C., Beaupoil-Hourdel, P. & Le Mené M., 2021, Children's socialization to multi-party interactive practices: Who talks to whom about what in family dinners, in L. Caronia (eds), *Language and Social Interaction at Home and in School*, Amsterdam: John Benjamins, p. 46-85.
- Navarrete-González A., 2021, The expression of contrast in Catalan Sign Language (LSC), *Glossa: a journal of general linguistics* 6(1), 29, p. 1-22.
- Nespor M. & Vogel I., 1986, *Prosodic phonology*, Dordrecht: Foris.
- Nespor M. & Sandler W., 1999, Prosody in Israeli Sign Language, *Language and Speech* 42, 2 & 3, p. 143-176.
- Parisse C., Blondel M., Caët S., Danet C., Vincent C. & Morgenstern A., 2022, Multidimensional Coding of Multimodal Language in Multi-Party Settings, *Proceedings of the 13th Conference on Language Resources and Evaluation* (LREC 2022), p. 2781-2787.
- Palancar E. L. & Vanhove M., 2021, Présentation générale : Des clivées et autres constructions focalisantes. *Faits de langues*, 52(1), p. 5-11.
- Pouw W., Trujillo J.P. & Dixon J.A., 2020, The quantification of gesture-speech synchrony: A tutorial and validation of multimodal data acquisition using device-based and video-based motion tracking. *Behavior Research Methods*, 52(2), p. 723-740. <https://doi.org/10.3758/s13428-019-01271-9>
- Ronveaux C. & Simon A.-C., 2005, La gestualité prosodique au service de l'objet enseigné, in *Conference proceedings Gestures and the organization of social*

- interaction: ethnomethodological and conversational perspectives.*
<http://hdl.handle.net/2078.1/83612>
- Rohrer, P. L., Florit-Pons, J., Vilà-Giménez, I., & Prieto, P. (2022). Children Use Non-referential Gestures in Narrative Speech to Mark Discourse Elements Which Update Common Ground, *Frontiers in Psychology*, 12, 661339. <https://doi.org/10.3389/fpsyg.2021.661339>
- Rossi M., 1999, *L'intonation, le système du français : description et modélisation*, Ophrys.
- Rossi M., 2011, L'intonation modale, *Modèles linguistiques* 63, p. 117-129.
- Roustan B. & Dohen M., 2010, Co-production of contrastive prosodic focus and manual gestures: temporal coordination and effects on the acoustic and articulatory correlates of focus, *5th International Conference on Speech Prosody* (Speech Prosody 2010), May, Chicago, United States 100110:1-4, p. 2010.
- Sandler W., 1999, Prosody in two natural language modalities *Language and Speech*, 42 2&3, p. 127-142.
- Sandler W. & Lillo-Martin D., 2005, *Sign language and linguistic universals*, Cambridge: Cambridge University Press.
- Séguinot A., 1977, L'accent d'insistance en français standard, in F. Carton, D. Hirst, A. Marschal & A. Séguinot (eds), *L'accent d'insistance : emphatic stress*, *Studia Phonetica* 12, p. 1-58.
- Shi R., Morgan J. & Allopenna P., 1998, Phonological and acoustic bases for earliest grammatical category assignment: a cross-linguistic perspective, *Journal of Child Language* 25, p. 169-201.
- Sloetjes H. & Wittenburg P., 2008, Annotation by category - ELAN and ISO DCR, in *Proceedings of the 6th International Conference on Language Resources and Evaluation* (LREC 2008).
- Stavropoulou P. & Spiliotopoulos D., 2011, Prosodic realization of topics, *19th International Symposium of Theoretical and Applied Linguistics*, p. 413-420. Thessaloniki, Greece: Department of Theoretical and Applied Linguistics, Aristotle University of Thessaloniki.
- Tang G., Brentari D., González C. & Sze F., 2010, Crosslinguistic variation in prosodic cues, in D. Brentari (eds), *Sign Languages*, Cambridge University Press, p. 519-542.
- Vallduví, E. & Vilkuna, M., 1998, On rheme and kontrast, in P. Culicover, & L. McNally (eds), *The limits of syntax*, New-York: Academic Press, p. 79-108.
- Wilbur R. B., 1994a, Stress, Focus and Extrametricality in ASL [LSA version January, 1994] [revised 4/94]
- Wilbur R. B., 1994b, Foregrounding structures in ASL. *Journal of Pragmatics. An interdisciplinary monthly of language studies* 22, p. 647-672.
- Wilbur R. B. Ms revised, 1997, Prosodic Structure of American Sign Language à paraître in Lynch (ed) *The cognitive Science of Prosody*. Amsterdam: Elsevier.
- Wilbur R. B., 1999, Stress in ASL: Empirical evidence and linguistic issues, *Language and Speech* 42, 2&3, p. 229-250.
- Wilbur R. B. & Malaia E., 2018, A new technique for analyzing narrative prosodic effects in sign languages using motion capture technology, in A. Hübl & M. Steinbach (eds), *Linguistic Foundations of Narration in Spoken and Sign Languages*, Amsterdam: John Benjamins, p. 15-40.

- Wilbur R. B. & Patschke C. G., 1998, Body leans and the marking of contrast in American Sign Language, *Journal of Pragmatics* 30, p. 275-303.
- Wilbur R. B. & Nolen S., 1986, The duration of syllables in American Sign Language, *Language and speech* 29, p. 263-280.
- Wilson D. & Wharton T., 2006, Relevance and prosody, *Journal of Pragmatics* 38(10), p. 1559-1579.
- Zenghui L., Shufen L. & Lei Z., 2020, The development of prosodic focus-marking and declarative question intonation in Thai learners' Mandarin, *10th International Conference on Speech Prosody*, 25-28 may, Tokyo, Japan.