



<https://bulletin.vals-asla.ch> · e-ISSN: 2813-7973 · p-ISSN: 1023-2044

Ursi, B. & André, V. (2021). Corpus: exploration, médiation et autonomisation. L'utilisation du concordancier de la plateforme FLEURON en cours de FLE. *Bulletin suisse de linguistique appliquée* (Numéro spécial 2), 129-148. <https://doi.org/10.26034/ne.vals-asla.2021.1269>

Cet article est publié sous la licence *Creative Commons Attribution 4.0 International* (CC BY):
<https://creativecommons.org/licenses/by/4.0>



© Biagio Ursi, Virginie André, 2021

Corpus: exploration, médiation et autonomisation.

L'utilisation du concordancier de la plateforme FLEURON en cours de FLE

Biagio URSI

Aix-Marseille Université & Laboratoire Parole et Langage – UMR 7309 (CNRS, AMU)
Département de Lettres modernes, Aix-Marseille Université
29, Avenue Robert Schuman, 13621 Aix-en-Provence Cedex 01, France
biagio.ursi@univ-amu.fr

Virginie ANDRÉ

ATILF – UMR 7118 (Université de Lorraine et CNRS)
44, Avenue de la Libération BP 30687, 54063 Nancy Cedex, France
virginie.andre@univ-lorraine.fr

In this paper, using a specific didactic approach, data-driven learning, we study the use of the FLEURON platform (<https://fleuron.atilf.fr/>) and, in particular, its concordancer by FLE (French as a foreign language) teachers and allophone students of different levels. Learners rely on this tool to observe data, identify specific forms, propose categorizations and generalize the use of these forms based on real occurrences and sufficiently broad contexts. The concordancer serves as a resource for autonomous learning which, after initial trial and error, results in a real linguistic "investigation". Based on classroom recordings, we propose a multimodal interactional analysis of sequences where the use of certain words in context is illustrated through the mobilization of conversational data within the FLEURON platform. This will lead us to identify the emerging features of the learning process among users of the platform, as for learners' initiatives and their responses to teacher's solicitations.

Keywords:

digital platforms, interactional corpora, French as a foreign language, concordancer, data-driven learning, multimodality, interaction analysis.

Mots-clés:

plateformes numériques, corpus de la langue parlée en interaction, FLE, concordancier, apprentissage sur corpus, multimodalité, analyse des interactions.

1. Introduction

L'exploitation des corpus à des fins didactiques s'inscrit dans la suite et en complément de l'utilisation des documents authentiques, défendue depuis les années 1970 (Abé et al. 1979). Néanmoins, de cette exploitation, qui peut prendre différentes formes, émergent de nouvelles méthodologies dans le champ de l'enseignement et de l'apprentissage des langues (Boulton & Tyne 2014), notamment du français langue étrangère (FLE) et des compétences interactionnelles (André 2019). Parmi les différentes façons d'exploiter des corpus, nous nous intéresserons à la démarche associée au *data-driven learning* (Johns 1991; Aston 2001) qui propose à des apprenant·e·s de langue

d'interroger des corpus à l'aide d'un concordancier, puis d'observer les résultats, de les analyser, de les classer et d'induire les règles de fonctionnement des faits langagiers recherchés.

Le dispositif FLEURON (Français Langue Étrangère Universitaire: Ressources et Outils Numériques, <https://fleuron.atilf.fr/>) propose à des étudiant·e·s étranger·e·s, qui souhaitent se préparer à interagir dans les situations de communication qu'elles·ils vont rencontrer en France, un corpus d'interactions outillé d'un concordancier. Le corpus multimodal de FLEURON est constitué d'interactions universitaires et de situations vécues par des étudiant·e·s en dehors du campus (achats de billets en gare, de pain dans une boulangerie, explications d'itinéraires, etc.). Dans le cadre de cette étude, nous nous intéressons aux activités et aux processus analytiques mis en œuvre par des apprenant·e·s qui focalisent leur attention sur les résultats de leurs recherches dans le concordancier de FLEURON. Ces recherches, par chaînes de caractères, concernent des mots isolés, des blocs de mots ou des constructions spécifiques en français parlé. Nous proposons une analyse multimodale (Traverso 2016; Mondada 2017) des interactions qui prennent place lors de l'utilisation du dispositif FLEURON en cours de langue, entre les apprenant·e·s et entre les apprenant·e·s et l'enseignant·e qui les accompagne. Nous présenterons tout d'abord le cadre de cette étude et la mise en place des expérimentations avant d'exposer leurs analyses.

2. L'apprentissage sur corpus et le domaine de l'interaction

Le *data-driven learning*, ou en français l'apprentissage sur corpus (ASC) (Boulton & Tyne 2014), a fait ses preuves dans le domaine de l'enseignement et de l'apprentissage de l'anglais écrit (Breyer 2011; Boulton & Cobb 2017). Cependant, il n'existe que de rares travaux mettant en œuvre la démarche de l'ASC avec des corpus oraux (voir notamment Zorzi 2001 pour l'italien, Mauranen 2004 pour l'anglais ou Kerr 2013 pour le français) et encore plus rares avec des corpus multimodaux (Braun 2007). Les premières expérimentations de l'ASC pour acquérir des compétences interactionnelles en français s'inscrivent dans le cadre du projet FLEURON qui a expérimenté les principes de l'ASC avec son corpus multimodal et son concordancier (André 2017; André & Ciekanski 2018). L'acquisition de compétences interactionnelles permet à un·e apprenant·e de mobiliser les ressources (lexicales, syntaxiques, mais aussi multimodales) appropriées pour réussir à participer à un échange finalisé et situé avec un·e ou plusieurs locuteur·trice·s. Ces compétences reposent sur la triple capacité pour un·e locuteur·trice de s'adapter à la situation de communication et à tous les éléments qui la composent et qui influencent les interactions, de s'ajuster aux propos de son interlocuteur·trice pour produire des actions séquentiellement adéquates, et de construire collaborativement un discours commun.

En suivant la démarche de l'ASC, les apprenant·e·s sont amené·e·s à changer de paradigme, traditionnellement centré sur la transmission des connaissances, et à entrer dans un processus d'apprentissage de la langue qui prend appui sur trois étapes: observation – catégorisation – généralisation (Johns 1991). Ainsi, tout d'abord, l'apprenant·e est amené·e à observer les données, extraites par un concordancier et consultables sous la forme de listes de lignes de concordances. Elle·il analyse la distribution des occurrences d'un mot (grammatical ou lexical), c'est-à-dire qu'elle·il compare ses conditions d'apparition et ses usages. Elle·il procède ensuite à une catégorisation de ces données et réunit les occurrences qui ont le même sens, qui ont le même cotexte, ou encore qui s'intègrent dans des *chunks* ou des blocs de mots relevant d'une analyse syntaxique. Enfin, en s'appuyant sur cette catégorisation, l'apprenant·e induit les règles de fonctionnement des occurrences recherchées. Les activités cognitives et métacognitives développées lors de ces trois étapes semblent favoriser l'appropriation de la langue, de ses usages et de son fonctionnement en contexte. L'apprenant·e devient actif·ve dans son apprentissage en prenant le rôle d'un·e apprenti·e-linguiste ou d'un "Sherlock Holmes" (Johns 1997: 101). Gavioli (2001) et Bernardini (2004) parlent de *learner as researcher* pour évoquer les activités dans lesquelles s'investissent les apprenant·e·s en analysant des lignes de concordances. L'enseignant·e change également de rôle. Chambers (2010) explique ce changement en ces termes:

The role of the teacher also changes fundamentally, as s/he is no longer the sole source of knowledge about target language, but rather a facilitator of the learning process, helping the learners to interpret the data, and giving them advice on how best to search the corpus and analyse their results. (Chambers 2010: 355)

L'enseignant·e devient un·e guide, un·e conseiller·e ou un·e accompagnateur·trice afin d'impulser la démarche de l'ASC et les activités associées. En outre, en apprenant comment apprendre, l'apprenant·e s'autonomise au fur et à mesure de ses expériences. À ce propos, Aston souligne: "[p]erhaps the greatest attraction of corpora for language pedagogy is their potential for autonomous learning" (2001: 41).

La linguistique de corpus associée à la recherche en didactique des langues a impulsé l'ASC. La méthodologie d'analyse des données en corpus a été exploitée par des apprenant·e·s de langue pour entrer dans une démarche active de compréhension et d'analyse des données. Cette méthodologie se révèle particulièrement intéressante et pertinente pour des apprenant·e·s qui souhaitent développer leurs compétences socio-interactionnelles à l'oral, dans la mesure où elles·ils apprennent à communiquer de manière efficace en situations réelles. Tout d'abord, les manuels de langue ne présentent pas ou peu le français parlé en interaction tel qu'il est réellement produit (Giroud & Surcouf 2016). Ils prescrivent un enseignement de la norme, souvent en référence à l'écrit, et s'éloignent de l'exposition à une langue réelle et située.

Ensuite, la description dans les documents de référence du français parlé en interaction est encore parcellaire (Traverso 2016: 13-14) et les enseignant·e·s de FLE peuvent être légitimement démuni·e·s pour proposer à leurs apprenant·e·s des règles de fonctionnement. Ainsi, l'ASC permet aux apprenant·e·s ainsi qu'aux enseignant·e·s d'observer des interactions situées et de saisir des phénomènes langagiers en usage.

3. Les expérimentations de FLEURON

Dans le cadre du projet FLEURON, nous avons souhaité documenter les usages que des apprenant·e·s de langue étrangère font de cette plateforme et du dispositif d'apprentissage qu'elle propose¹. Notre attention s'est focalisée sur le concordancier multimédia (André & Ciekanski 2018). Cet outil permet non seulement de rassembler et de lister les occurrences d'une même forme linguistique (ou d'une même chaîne de caractères) présente dans différentes transcriptions, mais aussi de visionner les extraits vidéo des ressources correspondantes. En d'autres termes, toutes les occurrences présentées en liste dans les résultats du concordancier sont des hyperliens qui permettent d'afficher la ressource dans un nouvel onglet et d'avoir ainsi accès au signal et à la transcription de l'enregistrement dans lequel l'occurrence est prononcée. Plus précisément, la vidéo s'ouvre au moment où l'occurrence sur laquelle l'apprenant·e a cliqué est prononcée. Le titre de l'extrait vidéo ainsi qu'une description succincte de la situation de communication et des thèmes qui y sont abordés fournissent également des éléments de contextualisation précieux pour l'exploration de la ressource.

Nos expérimentations se sont déroulées avec des apprenant·e·s du Département de Français Langue Étrangère (DéFLE) au Centre de Langues Yves Chalon (CLYC) de l'Université de Lorraine avec la collaboration de Florence Poncet, enseignante de FLE. Les données sur lesquelles ce travail se fonde sont issues d'une campagne de collecte menée au CLYC en septembre 2019, pendant les cours intensifs de langue française destinés aux étudiant·e·s internationaux·ales, dans le cadre du programme ERASMUS. Nous avons enregistré les séances d'étudiant·e·s face à un ordinateur, seul·e·s et en binômes, accompagné·e·s d'un enseignant/tuteur qui a surtout assuré la médiation des fonctionnalités du site internet et des spécificités de l'interface numérique, auxquelles les apprenant·e·s avaient déjà été initié·e·s. Pour cet article, nous nous focaliserons uniquement sur une seule des configurations expérimentées, celle avec deux étudiant·e·s et un enseignant. Pendant ces expériences, les apprenant·e·s ont exploré les ressources de FLEURON en s'appuyant sur le concordancier; en particulier, elles·ils ont recherché des

¹ Cette étude expérimentale est soutenue par le projet OLKi (*Open Language and Knowledge for Citizens*) de Lorraine Université d'Excellence (LUE), ANR-15-IDEX-04-LUE.

formes linguistiques qu'elles·ils ont, dans la plupart des cas, elles·eux-mêmes choisies.

Concernant le dispositif de captation vidéo, nous avons utilisé deux caméras: l'une focalisée sur les participant·e·s, qui a offert une prise de vue latérale et générale sur les postures des corps et les orientations mutuelles des participant·e·s; l'autre focalisée sur l'écran. Nous avons opté pour cette dernière option, afin d'avoir accès aux détails des pointages et des gestes accomplis devant l'écran par les participant·e·s, plutôt que d'effectuer une captation depuis l'ordinateur (qui aurait permis de documenter exclusivement les mouvements du curseur et la transition entre les différentes fenêtres de la plateforme à l'écran). Par la suite, nous avons édité les données afin d'obtenir dans un même fichier les vidéos synchronisées des deux prises de vue, comme nous pouvons le voir dans la Fig. 1. Ici, des apprenant·e·s travaillent en binôme face à un ordinateur, en présence d'un enseignant; elles·ils parcourent une liste d'occurrences du concordancier FLEURON (qui apparaît dans la prise de vue incrustée en haut à droite, voir 4.1 pour des détails ultérieurs).

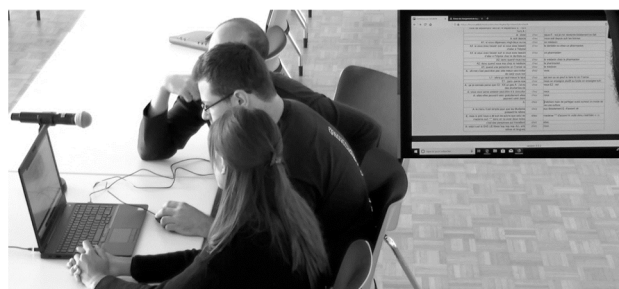


Fig. 1: Double captation vidéo

4. Analyse

L'analyse séquentielle et multimodale de trois extraits représentatifs nous permet de souligner des éléments de réflexion concernant les modalités d'exploration de la base FLEURON par le biais du concordancier et, en particulier, la découverte chez les apprenant·e·s des régularités linguistiques du français parlé à partir de corpus conversationnels. La nature de cette étude est donc principalement qualitative, mais nous avons observé des éléments récurrents qui se prêteraient à une appréciation quantitative. Notre regard analytique est focalisé sur le rôle des ressources multimodales mobilisées par les apprenant·e·s, dans le cadre d'une activité ciblée qui s'inspire des principes de l'apprentissage sur corpus et qui se fonde sur des documents conversationnels authentiques, en vue du développement des compétences linguistiques des apprenant·e·s. Comme nous le verrons, même l'initiative de l'apprenant·e à partir de problèmes locaux peut amener à un processus de découverte quant aux spécificités de la langue parlée en interaction et de

certaines formes linguistiques, notamment la préposition "depuis" (4.1) et la locution verbale "valoir mieux" (4.2).

4.1 Observation des formes linguistiques et problèmes émergents

Dans l'extrait suivant, trois participant·e·s sont présent·e·s. Deux étudiant·e·s Erasmus en troisième année de licence ayant un niveau de compétence avancé en français², LAU (Laura), une étudiante italienne, et MIR (Mircea), un étudiant roumain, se penchent sur la liste des occurrences de la préposition "depuis", qu'elles·ils ont elles·eux-mêmes choisi de rechercher; Georges (GEO) est l'enseignant de français. Cette liste est affichée dans le concordancier et elle est parcourue verbalement par les deux apprenant·e·s au début de l'extrait 1.

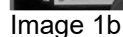
(1) CLYC_Groupe 1_Depuis_a (18:48)

```

01  LAU  cinq ans/ (0.3) oui (.) °ça ça indique toujours le [temps°
02  MIR                                     [depuis
03      cinq ans
04  GEO  depuis cinq ans/
05      (0.3)
06  MIR  x xx
07      (4.0)
08  MIR  hm la première année/
09      (0.9)
10  GEO  depuis la première année
11      (0.3)
12  MIR  dix minu:tes depuis quelque euh année/
13      (0.5)
14  MIR  *euh depuis: quelle année/
15  mir  *secouements de tête--->
16      (0.3) * (0.7)
17  mir  -->*
18  MIR  euh: depuis l'année dernière (0.4) de nouveau du temps
19      (0.5)
20  MIR  xxx
21      (2.5) * (0.3)
22  lau  *....>
23      *....>
24  LAU  et ici/*#1
25  lau  >...*parcourt du doigt une ligne du concordancier-->
26      >...*se penche vers l'écran

```

² B1/B2, d'après le test effectué par l'enseignante responsable du module de formation selon les critères du *Cadre européen commun de référence pour les langues*. La plateforme FLEURON a été précédemment présentée aux apprenant·e·s qui ont participé à nos expériences.

Image 1c[illegible]

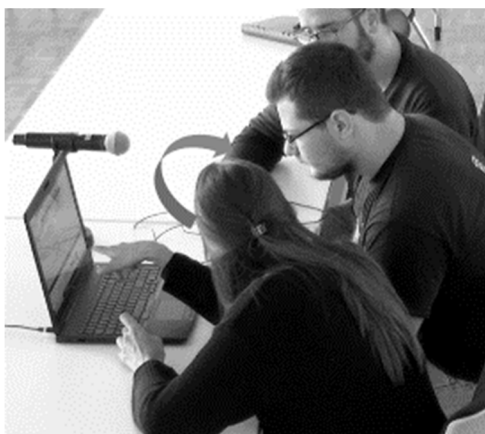


Image 2

24 (0.5)
 25 GEO alors on+ va VOIR*+ on peut* voir
 lau >,,*
 -->+,,,,,,,,,+ *se redresse->
 26 (0.8)*(0.6)
 lau -->*
 27 MIR ((clicque sur l'occurrence du concordancier))

Dès les premières lignes, les deux apprenant·e·s sont focalisé·e·s sur le contexte droit des quatorze occurrences de *depuis* listées dans le concordancier (image 1c). La lecture des différents éléments qui suivent cette forme linguistique ne s'inscrit pas dans une simple activité de restitution, mais dans une véritable démarche réflexive à partir de données linguistiques en contexte qui est réalisée de manière récurrente dans nos expériences. De ce point de vue, le tour ligne 1 est représentatif dans son articulation tripartite: la lecture du syntagme introduit par *depuis* ("cinq ans") est suivie d'une pause et, surtout, d'une particule ("oui") qui signale l'enregistrement de cette unité par rapport à sa simple mention. Une généralisation est alors produite de manière située, en guise de confirmation d'une hypothèse déjà formulée ("ça ça indique toujours le temps"). La structure de cette généralisation permet de relier la singularité de l'occurrence à une constatation qui est corroborée par la pluralité de données à disposition de l'apprenante-enquêtrice, par le biais d'une dislocation à gauche et d'une caractérisation fréquentielle ("toujours"), respectivement. Cette généralisation est prononcée à voix basse et, pour cette raison, ne fait pas l'objet d'une validation de la part de l'enseignant.

De son côté, Mircea est engagé dans une activité spécifique: il parcourt les occurrences affichées dans le concordancier. Il lit parfois les mots qui suivent la forme cible (par ex. "la première année", l. 08) parfois la totalité du syntagme prépositionnel (par ex. "depuis cinq ans", l. 02-03). Les mentions des lignes du concordancier sont hétéro-répétées par Georges, qui alimente ainsi la progression de l'activité par la reprise de formes étendues; ces ressources jouent donc un rôle structurant pour l'activité en cours (cf. Berger 2008). Lignes

01-14, l'intonation énumérative finale des interventions permet d'inscrire ces tours de parole dans une même trajectoire actionnelle.

Au sein de l'activité de lecture, des éléments problématiques émergent. Mircea revient sur la production de la ligne 12 "depuis quelque euh année", où un coup de glotte est perceptible entre les deux mots qui suivent la forme cible, et la répare par sa forme correcte "euh depuis: quelle année" (l. 14). Cette auto-réparation est accompagnée d'un secouement de tête évident chez l'apprenant, comme le montre la transcription multimodale. Comme Laura, Mircea ajoute lui aussi une caractérisation qui permet de spécifier rétrospectivement les mots situés à droite de la forme cible, qui font l'objet de sa lecture, et verbalise une généralisation en train de se faire ("de nouveau du temps", l. 16). Ligne 20, un autre problème est signalé par Laura, ce qui va déclencher une longue séquence de réparation (Schegloff et al. 1977; Schegloff 2007) présentant des réarrangements corporels notables. Elle se penche vers l'écran (image 1a) et indique la zone occupée par la transcription problématique (image 1b), à l'aide d'un pointage dynamique. En effet, elle parcourt du doigt la ligne du concordancier et puis s'arrête pour lire ce qui suit le *depuis* ("depuis le: (0.6)"), ou, pour mieux dire, afin de signaler un manque notable ("°après° c'est fini", l. 23), le pointage est alors figé (image 2). Pendant la lecture, elle réoriente son corps se penchant davantage en avant et tournant sa tête vers sa droite pour chercher un contact visuel avec l'enseignant, qui est ainsi interpellé: une reconfiguration du cadre participatif (Goodwin & Goodwin 2004) a lieu par rapport à la position initiale des participant·e·s, où leurs têtes étaient alignées (comme le montre la Fig. 1) et leurs regards focalisés sur l'écran.

Ligne 25, Georges réagit en formulant une instruction indirecte, qui signale le caractère interactif de l'interface du concordancier et évoque la possibilité de visionner l'enregistrement correspondant à la ligne problématique (une fonctionnalité proposée par le concordancier multimédia). Cette instruction est présentée comme envisageant une opération qui mobilise la totalité des participant·e·s (voir notamment l'utilisation du pronom sujet "on", dans les deux expressions verbales de ce tour), en tant qu'observateur·trice·s face aux données. C'est Mircea, la souris dans la main, qui accomplit l'action instruite, c'est-à-dire l'ouverture de la fenêtre présentant la vidéo obtenue par un clic sur l'occurrence désignée (l. 27). Ainsi, la recherche de l'occurrence problématique est réalisée dans le contexte élargi de la transcription intégrale de l'enregistrement (le deuxième extrait constitue la suite du premier).

(2) CLYC_Groupe 1_Depuis_b (18:48)

```

28   MIR   <(8.2) ((fait défiler la transcription vers le bas))>
29   MIR   ah non
30       <(7.0) ((fait défiler la transcription vers le haut))>
31   LAU   *[là]*##3
32   MIR   *[hm]* ((petit rire))*
      lau   *....*pointe l'écran--*,,>

```

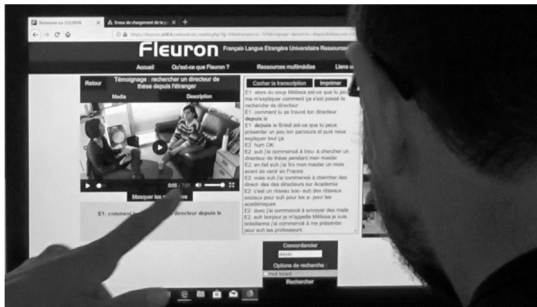


Image 3

33 (0.4) * (2.9)
 lau >, , , *
 34 MIR ((clique sur play))
 35 vid <((audio de la vidéo en ligne)) comment tu as trouvé ton
 36 directeur depuis le:: depuis le brésil est-ce que tu peux
 37 présenter un peu ton [parcours et puis [nous ex[pliquer&
 38 MIR [ah::
 39 LAU [ah o[kay
 40 MIR [c'est ça
 41 vid &tout ça hum okay\>
 42 MIR ((clique sur pause))
 43 GEO alors voilà\
 44 MIR depuis le brésil c- c'était ça
 45 (0.6)
 46 LAU mh mh
 47 (0.3)
 48 MIR mh mh
 49 (1.0)
 50 GEO donc c'é*tait un:*#4 quelque chose qui [arri-
 lau *.....*se penche en avant, se tourne vers GEO
 51 LAU [elle a *ré*pété*#5
 lau *..*-----*, , , >

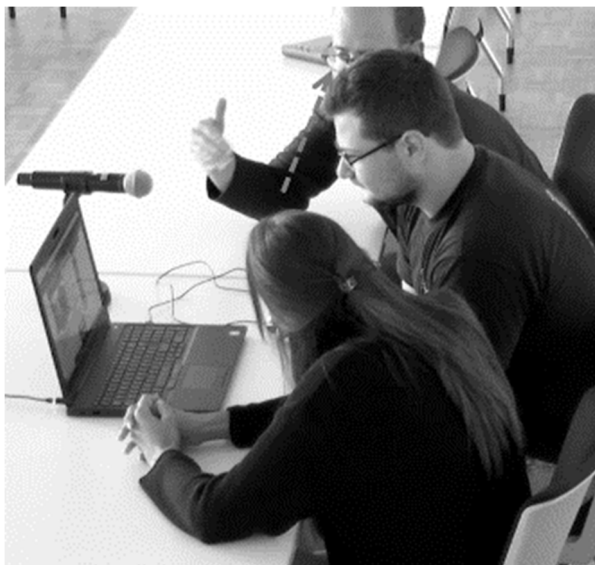


Image 4



Image 5

52 LAU e*lle a *con*tinué*#6 a*près
 lau >, * *..*-----*, , , *(pointe l'écran)



Image 6

53 GEO voilà

Après l'affichage de la transcription de l'extrait intitulé "Témoignage: rechercher un directeur de thèse depuis l'étranger", la recherche de l'occurrence problématique prend un peu de temps (voir les longs silences lignes 28 et 30, dûs au défilement du texte de la transcription) et comporte également des erreurs (l. 29 Mircea s'arrête sur un tour de parole qui n'est pas celui recherché et s'exclame par un "ah non"³), mais cette activité est une activité collective, à laquelle les deux apprenant·e·s contribuent de manière active. Mircea manipule la souris et fait défiler le texte de la transcription jusqu'au moment où Laura indique l'expression recherchée de manière multimodale, par un pointage (image 3) et un déictique spatial ("là", l. 31). En même temps, l'étudiant prononce un "hm" et un petit rire qui signalent également l'aboutissement de sa recherche, il n'a pas encore posé le curseur sur le tour verbal ciblé mais il a repéré visuellement l'occurrence problématique, simultanément à Laura. Mircea lance alors la vidéo. Les lignes 35, 36, 37 et 41 reproduisent la transcription du tour de parole prononcé par la locutrice qui interviewe la doctorante brésilienne sur la recherche de son directeur de thèse en France.

Suite au visionnage et, en particulier, à l'écoute des données, Mircea et Laura reconnaissent finalement l'emplacement séquentiel de la production inachevée qui a posé problème ("depuis le:", l. 23, ex. 1): il s'agit d'une auto-réparation avec un premier syntagme non finalisé ("depuis le::"), suivi de sa réalisation complète ("depuis le Brésil", l. 36). La compréhension des deux apprenant·e·s est affichée par des *change-of-state tokens* (Heritage 1984), exprimant l'enregistrement d'une nouvelle information, et des marques de réception ("ah okay", "ah:: c'est ça", l. 38-40).

Mircea arrête l'enregistrement (l. 42), l'enseignant projette la clôture de cette séquence ("alors voilà", l. 43) et l'étudiant verbalise le syntagme complet, qui

³ Ce tour de parole verbalise ainsi une auto-réparation.

représente l'auto-réparation de l'occurrence problématique dans la transcription ("depuis le Brésil c- c'était ça", l. 44). Les apprenant·e·s s'orientent donc vers une prise en compte de cette résolution (l. 46-48), mais c'est Georges qui essaie de proposer une explication au phénomène observé et la relie à ce qui vient d'être dit ("donc c'était un: quelque chose qui arri-", l. 50). C'est à ce moment-là que Laura se penche en avant, encore une fois pour établir un contact visuel avec l'enseignant (images 4 et 5), et donne une explication de ce qui a été remarqué ("elle a répété elle a continué après", l. 51-52). Son discours est rétrospectivement lié à l'observation effectuée sur l'écran, que l'apprenante pointe à deux reprises pendant son intervention (image 6). Enfin, l'enseignant valide cette explication et produit le marqueur de clôture "voilà" (l. 53) qui permet de fermer la séquence initiée par Laura à la ligne 20 de l'extrait 1.

Dans cette section, nous nous sommes focalisé·e·s sur le caractère émergent de l'observation de données à partir d'un concordancier multimédia. Dans la première partie de l'extrait 1, l'activité de lecture est accomplie de manière collaborative par les participant·e·s et est ponctuée d'observations généralisantes. Le caractère indéterminé d'une occurrence, observé par une apprenante, a donné lieu à une séquence de réparation étendue (Schegloff 2007). À l'aide de deux extraits représentatifs, nous avons pu montrer que les initiatives des étudiant·e·s peuvent porter sur des phénomènes locaux et très courants à l'oral tels qu'une auto-réparation (ce qui est réalisé par l'intervieweuse dans l'enregistrement FLEURON visionné par les apprenant·e·s, ex. 2). Ces initiatives sont caractérisées par une reconfiguration du cadre participatif (Goodwin & Goodwin 2004) et une réorientation du corps de l'apprenant·e, qui se trouve dans une position contrainte face à l'écran de l'ordinateur. En particulier, la recherche d'un contact visuel avec l'enseignant marque la mise en évidence du problème et sa résolution par l'apprenant·e, qui, dans ce dernier cas, sollicite la validation du participant "expert". La mobilisation des ressources multimodales (notamment, gestes de pointage dynamiques ou figés et reconfigurations corporelles) joue un rôle primordial dans les étapes saillantes (repérage et explication) portant sur le tour problématique. Son caractère problématique est, en définitive, associé à la transcription non suffisamment contextualisée, telle qu'elle apparaît dans la colonne droite du concordancier. Par ailleurs, la transcription d'un phénomène courant en français parlé tel qu'une auto-réparation et sa visibilité dans le signal correspondant permettent aux apprenant·e·s de se familiariser avec ce type de phénomènes, de les produire et de les interpréter comme des faits langagiers qu'il est possible de mobiliser dans la langue parlée en interaction.

4.2 La focalisation sur les formes linguistiques

L'extrait suivant montre un cas récurrent dans l'observation des occurrences d'une même forme linguistique, toujours sélectionnée par les apprenant·e·s, où la focalisation sur des éléments spécifiques est guidée par l'enseignant. Dans

l'extrait suivant, deux étudiantes Erasmus italiennes⁴, Maria et Elena, ont choisi de focaliser leur attention sur la forme *mieux* et elles remarquent plusieurs attestations de la collocation *vaut mieux*. Comme pour la situation précédente, cette activité se déroule en présence de Georges, l'enseignant.

(3) CLYC_Groupe 2_Vaut mieux (14:39)

01 GEO et donc euh et là qu'est-ce que vous >par exemple<
02 si on se focalise#7 sur VAUT MIEUX



Image 7

03 *(.) qu'est-ce que vous remarquez*#8
mar *penche sa tête en avant-----*



Image 8

04 (0.7)
05 GEO au niveau du pronom
06 (3.7)
07 GEO >donc< f'c'est-à-fdire mieuxf qui estf précédé par vaut/f#9
geo fbuste en arrière-----f
f.....fmain/table f.....fm -> g

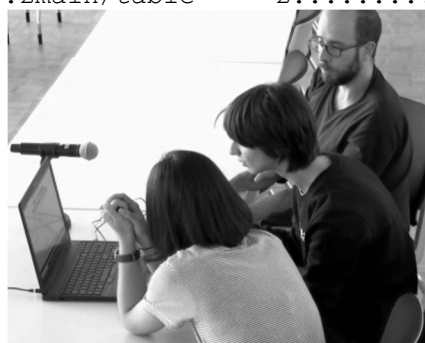


Image 9

⁴ D'après l'enseignante responsable du module de formation, les apprenant·e·s ont en français le même niveau de compétence que les apprenants de l'extrait 1.

08 (0.1)£(0.6)£ et vaut/£ (0.2) il est précédé par un pronom/
geo m/t £,,,,,£ £buste en avant-->

09 (0.4)

10 ELE euh£: [oui]

11 MAR [oui]

geo ->£(buste en avant)

12 GEO £c'est £[il vaut oui]#10

13 ELE [il vaut mieux] oui [c'est IM]personnel

geo £.....£tête en avant

14 GEO [°il vaut°]

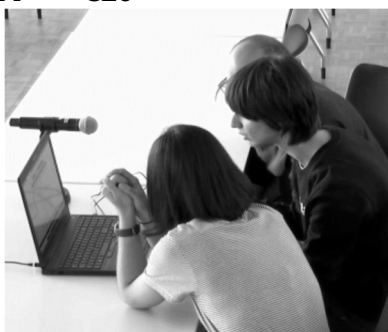


Image 10

```

15  MAR    il vaut mieux\
16  <(2.0) ((MAR pose le curseur sur une occurrence, le déplace))>
17  GEO    tou*jours/ (.) on va voir fon* va:f vous pouvez cli[quer/
mar      *.....*redresse sa tête
geo      fmvt têtef
18  MAR
19  sur une occurrence))

```

Au début de cet extrait, Georges sollicite les participantes pour qu'elles prêtent attention aux attestations affichées dans le concordancier et, en particulier, aux occurrences de *vaut mieux* (cette expression linguistique est prononcée de manière saillante l. 02). L'enseignant répare son tour initiatif qui est ainsi articulé en deux moments: (i) une *si*-proposition qui présente l'objet attentionnel sur lequel les participantes sont invitées à se focaliser ("si on se focalise sur VAUT MIEUX", l. 02); (ii) une interrogative partielle qui thématise la tâche observationnelle de l'activité ("qu'est-ce que vous remarquez", l. 03). De manière intéressante, les ressources grammaticales retracent la bipartition sur le plan pragmatique: si dans la première partie de ce tour multi-unités (Sacks et al. 1974) Georges utilise une marque pronominale inclusive ou générique ("on"), dans la deuxième partie, celle qui présente la forme d'une interrogative et projette donc une réponse, il s'adresse explicitement aux apprenantes par un pronom de deuxième personne pluriel ("vous"). Deux types de réponse sont fournies en réaction à ces deux unités.

La réponse des deux apprenantes à la première unité se manifeste de manière évidente sur le plan multimodal. Suite à la sollicitation attentionnelle produite par l'enseignant (via la *si*-proposition), Maria et Elena se penchent vers l'écran (le mouvement est particulièrement visible chez Maria, si on regarde l'image 8 et on la compare à l'image 7). Cette réponse corporelle, qui incarne la

focalisation attentionnelle précédemment sollicitée par l'enseignant, est un phénomène récurrent dans les expérimentations que nous avons pu enregistrer.

Ligne 5, Georges produit un incrément (Schegloff 1996; Ford et al. 2002) à sa question précédente, ce qui permet de mieux la spécifier. Cette extension mentionne la catégorie grammaticale que vise sa question ("au niveau du pronom"). Ensuite, il cherche à contextualiser sa question en retraçant l'historique des observations qui jusque-là ont été faites par les apprenantes ("c'est-à-dire mieux qui est précédé par vaut", l. 07) et il formule une véritable question portant sur la forme verbale *vaut*, qui est disloquée ("et vaut (0.2) il est précédé par un pronom", l. 08). Cette évocation du passé est accompagnée d'un réagencement multimodal: Georges éloigne son corps de l'écran (image 9). Les deux étudiantes produisent une réponse chorale ("oui" en chevauchement, l. 10-11) à la question de l'enseignant. Ce dernier se situe dans une position ambivalente: il répond lui aussi à sa propre question après avoir lu une occurrence du concordancier ("il vaut oui", l. 12), en se penchant vers l'écran (image 10).

Ligne 13, de son côté, Elena fournit une réponse plus complète: après avoir lu une occurrence ("il vaut mieux"), elle reproduit sa réponse positive ("oui") et donne une caractérisation morphosyntaxique précise du pronom en question ("c'est IMpersonnel"). Maria participe à cette activité suivant une temporalité distincte, après l'intervention de sa camarade, elle se limite à lire l'occurrence ("il vaut mieux", l. 15). Georges relance les participantes en questionnant la réalisation systématique du pronom ("toujours/", l. 17) et, pour inviter au visionnage de l'extrait, il mobilise la même structure que nous avons rencontrée dans le premier extrait analysé en 4.1 ("on va voir on va:", l. 17). Ensuite, il donne une instruction plus directe et s'adresse explicitement à Maria ("vous pouvez cliquer", l. 17): à ce moment-là, elle a positionné le curseur sur une occurrence qui n'est pas précédée d'un pronom. L'action instruite qui suit, comme dans l'extrait 1, consiste en un clic sur l'occurrence pour ouvrir la fenêtre qui permettra de visionner la vidéo correspondante (l. 18-19).

Dans les deux premiers extraits, la focalisation sur une ligne du concordancier émerge au cours de la lecture des occurrences et, en particulier, de ce qui se trouve "à droite du mot" recherché. Cette focalisation est opérée par l'un·e des apprenant·e·s. En revanche, dans ce troisième extrait, nous nous sommes penché·e·s sur une focalisation sollicitée par l'enseignant, qui essaie de diriger l'attention des apprenantes sur les occurrences affichées dans le concordancier et, plus précisément, sur ce qui se trouve ou ne se trouve pas "à gauche du mot" recherché (en l'occurrence, le pronom sujet de la forme verbale "vaut mieux").

5. Conclusion

Dans cette contribution, nous avons proposé une étude séquentielle et multimodale de trois extraits issus de deux séances de travail avec la plateforme FLEURON. Les analyses présentées dans la section 4 ont mis en évidence l'importance des ressources multimodales mobilisées chez les apprenant·e·s pendant ces séances. En particulier, le pointage et la réorientation du corps permettent d'initier des séquences de réparation qui visent à résoudre des problèmes émergents (section 4.1). Dans les extraits 1 et 2, le problème concerne un mode de production typique de la langue parlée, celui de l'auto-réparation. Par exemple, le regard de Laura se pose sur une ligne qui présente un problème pour elle: elle n'arrive pas à trouver le mot à la droite de la forme cible recherchée et remarque une absence notable dans ce contexte ("après c'est fini", l. 23, ex. 1), ce qui est démenti par le visionnage de la vidéo. Ce sont les va-et-vient entre la transcription ponctuelle telle qu'affichée dans le concordancier, la transcription complète et la vidéo de l'enregistrement correspondant qui peuvent fournir des éléments de contextualisation et, donc, d'explication des occurrences observées. La navigation entre les différents outils de la plateforme (concordancier, ressources, description du contexte) est possible grâce à la médiation de l'enseignant. Les marques linguistiques, les ressources multimodales et les manipulations au sein de la plateforme s'inscrivent dans une activité menée collaborativement par les binômes d'apprenant·e·s. Même si l'un des deux reste plutôt silencieux (en l'occurrence, Mircea dans l'extrait 1), il intervient avec les manipulations de la souris et avec des productions verbales ponctuelles, signalant sa compréhension effective de ce qu'il se passe dans l'interaction enregistrée.

Le court extrait de la section 4.2 nous a permis de présenter d'autres éléments récurrents dans notre corpus. Si, dans les deux premiers exemples, l'élément qui fait l'objet d'une enquête linguistique est repéré par une apprenante, dans le troisième c'est l'enseignant qui canalise les regards des deux participantes. En particulier, la sollicitation attentionnelle demandée par l'enseignant passe, sur le plan pragmatique, par une lamination actionnelle et, sur le plan de la structuration de la parole, par la mobilisation de ressources linguistiques s'inscrivant dans un tour multi-unités. La focalisation est accomplie de manière incarnée par les apprenantes⁵: elles se penchent vers l'écran et commencent la phase d'observation de données proprement dite, suite aux indications fournies par l'enseignant/médiateur. Dans l'extrait 3, l'initiative de l'apprenante est toujours présente, même si elle est silencieuse (en l'occurrence, Maria pose le curseur sur une attestation qui pose problème par rapport à la marque pronominale signalée par Georges et caractérisée par Elena). L'enseignant joue par moments le rôle de médiateur et sollicite la navigation entre les

⁵ Voir Mondada (2015) pour une analyse séquentielle détaillée dans des interactions en face-à-face, sans ordinateur.

ressources disponibles sur la plateforme, tout en conservant l'esprit collaboratif de l'activité en cours.

Les extraits analysés ici ont également fourni des exemples concernant la structuration de tours qui présentent des catégorisations spécifiant les traits sémantiques ("ça ça indique toujours le temps" pour *depuis*) ou grammaticaux ("c'est impersonnel" pour *il vaut mieux*) des formes cibles, ce qui illustre davantage l'usage acquisitionnel du concordancier. En particulier, les expériences que nous avons documentées attestent d'une structure tripartite composée des éléments suivants: (i) une occurrence et/ou son contexte, (ii) une particule qui est rétrospectivement liée à l'opération de lecture effectuée ("oui"), (iii) la catégorisation proposée (voir notamment le tour de Laura, l. 01, extrait 1; le tour de Elena, l. 13, extrait 3). La réalisation de la particule en position médiane n'est pas systématique. Quoi qu'il en soit, la catégorisation proposée par l'apprenant·e est le plus souvent ancrée dans une attestation (et s'inscrit donc dans un tour multi-unités), très rarement isolée.

Ce travail est exploratoire. Néanmoins, l'analyse multimodale des interactions accompagnant l'utilisation du dispositif FLEURON permet de rendre compte des activités cognitives et métacognitives réalisées par les apprenant·e·s, ainsi que du processus d'acquisition impulsé par ces activités. Parmi les ressources multimodales, les changements de postures par rapport à l'écran, l'orientation des regards ainsi que les pointages réalisés par les apprenant·e·s représentent des indices utiles pour le repérage de phases de "découverte", qui se distinguent de l'observation et de la simple lecture des lignes du concordancier. Les ressources linguistiques utilisées par l'enseignant, notamment les reprises (ex. 1) et les tours multi-unités (ex. 3), permettent de stabiliser les informations acquises pendant l'observation et de naviguer entre différents outils et onglets dans la plateforme numérique.

Nous avons mené des expérimentations avec des apprenant·e·s de FLE et, dans cette contribution, nous avons considéré une configuration participationnelle spécifique (celle d'un binôme d'étudiant·e·s universitaires ayant le même niveau de compétence linguistique, accompagné par un enseignant/tuteur) et nous nous sommes appuyé·e·s sur un dispositif d'enregistrement spécifique. Des études ultérieures pourront illustrer davantage le caractère émergent des observations et des commentaires (méta)linguistiques que les apprenant·e·s proposent pour les phénomènes langagiers accessibles depuis le concordancier multimédia. Une attention particulière a été accordée à la mobilisation, ici médiée par l'enseignant, des hyperliens donnant accès aux transcriptions complètes et aux enregistrements. Dans nos données, cette navigation s'avère centrale pour l'acquisition de certaines caractérisations proposées pour des phénomènes spécifiques, qui peuvent apparaître comme des récurrences ou des formes remarquables, voire problématiques, en parcourant les lignes du concordancier.

Par ailleurs, d'autres expérimentations sont en cours avec des apprenant·e·s adultes ayant différents profils en termes de niveau de scolarisation, niveau de langue, type de formation, compétences métalinguistiques, etc. dans différents contextes de formation (en autonomie, en classe). Elles permettront d'analyser également les effets de l'ASC sur le long terme et son efficacité quant à la rétention du fonctionnement de la langue en situation.

BIBLIOGRAPHIE

- Abé, D., Carton, F., Cembalo, S. M. & Régent, O. (1979). Didactique et authentique: du document à la pédagogie. *Mélanges pédagogiques*, 10, 1-14.
- André V. (2017). Un corpus multimédia pour apprendre à interagir en situations universitaires en France. In *Actes du troisième colloque international de l'ATPF "Enseigner le français: s'engager et innover"*, 292-315.*
- André, V. (2019). Des corpus oraux et multimodaux authentiques pour acquérir des compétences sociolangagières. In L. Gajo, J.-M. Luscher, I. Racine & F. Zay (éds.), *Variation, plurilinguisme et évaluation en FLE* (pp. 209-233). Berne: Peter Lang.
- André, V. & Ciekanski, M. (2018). Apprendre à interagir à l'oral à partir d'un concordancier multimodal: effets sur le développement de la conscience langagière et sur l'autonomie de l'apprenant dans le dispositif FLEURON. In C. Dejean, F. Mangenot, E. Nissen & T. Soubrié (éds.), *Actes du colloque EPAL 2018 (Échanger pour apprendre en ligne)*, Université Grenoble Alpes, 7-9 juin 2018.*
- Aston, G. (2001). Learning with corpora: An overview. In G. Aston (éd.), *Learning with corpora* (pp. 7-46). Houston: Athelstan.
- Berger, E. (2008). La reprise comme ressource interactionnelle en langue seconde. *TRANEL (Travaux neuchâtelois de Linguistique)*, 48, 43-61.
- Bernardini, S. (2004). Corpora in the classroom: An overview and some reflections on future developments. In J. Sinclair (éd.), *How to use corpora in language teaching* (pp. 15-36). Amsterdam/Philadelphia: John Benjamins.
- Boulton, A. & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, 67(2), 348-393.
- Boulton, A. & Tyne, H. (2014). *Des documents authentiques aux corpus. Démarches pour l'apprentissage des langues*. Paris: Didier.
- Braun, S. (2007). Integrating corpus work into secondary education: From data-driven learning to needs-driven corpora. *ReCALL*, 19(3), 307-328.
- Breyer, Y. A. (2011). *Corpora in language teaching and learning. Potential, evaluation, challenges*. Frankfurt am Main: Peter Lang.
- Chambers, A. (2010). What is data-driven learning? In A. O'Keeffe & M. McCarthy (éds.), *The Routledge handbook of corpus linguistics* (pp. 345-358). New York: Routledge.
- Ford, C. E., Fox, B. A. & Thompson, S. A. (2002). Constituency and the grammar of turn increments. In C. E. Ford, B. A. Fox & S. A. Thompson (éds.), *The language of turn and sequence* (pp. 14-38). Oxford/New York: Oxford University Press.
- Gavioli, L. (2001). The learner as researcher: Introducing corpus concordancing in the classroom. In G. Aston (éd.), *Learning with corpora* (pp. 108-137). Houston: Athelstan.

- Giroud, A. & Surcouf, C. (2016). De "Pierre, combien de membres avez-vous ?" à "Nous nous appelons Marc et Christian": réflexions autour de l'authenticité dans les documents oraux des manuels de FLE pour débutants. *SHS Web of Conferences* 27 - CMLF 2016.*
- Goodwin, C. & Goodwin, M. H. (2004). Participation. In A. Duranti (éd.), *A companion to linguistic anthropology* (pp. 222-244). Malden, MA: Blackwell.
- Heritage, J. (1984). A change-of-state token and aspects of its sequential placement. In J. M. Maxwell Atkinson & J. Heritage (éds.), *Structures of social action. Studies in conversation analysis* (pp. 299-345). Cambridge: Cambridge University Press.
- Johns, T. (1991). Should you be persuaded: Two samples of data-driven learning materials. *English Language Research Journal*, 4, 1-16.
- Johns, T. (1997). Contexts: The background, development and trialling of a concordance-based CALL program. In A. Wichmann, S. Fligelstone, T. McEnery & G. Knowles (éds.), *Teaching and language corpora* (pp. 100-115). London: Longman.
- Kerr, B. (2013). Grammatical description and classroom application. Theory and practice in data-driven learning. *Bulletin suisse de linguistique appliquée*, 97, 17-39.
- Mauranen, A. (2004). Spoken corpus for an ordinary learner. In J. Sinclair (éd.), *How to use corpora in language teaching* (pp. 89-105). Amsterdam/Philadelphia: John Benjamins.
- Mondada, L. (2015). Introduction de nouveaux référents et gestion de l'attention conjointe: les apports de l'analyse interactionnelle multimodale. In X. Gradoux, J. Jacquin & F. Merminod (éds.), *Agir dans la diversité des langues. Mélanges en l'honneur d'Anne-Claude Berthoud* (pp. 121-137). Bruxelles: De Boeck-Duculot.
- Mondada, L. (2017). Le défi de la multimodalité en interaction. *Revue française de linguistique appliquée*, XXII, 71-87.
- Sacks, H., Schegloff, E. A. & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696-735.
- Schegloff, E. A. (1996). Turn organization: One intersection of grammar and interaction. In E. Ochs, E. A. Schegloff & S. A. Thompson (éds.), *Interaction and grammar* (pp. 52-133). Cambridge: Cambridge University Press.
- Schegloff, E. A. (2007). *Sequence organization. A primer in conversation analysis*. Cambridge: Cambridge University Press.
- Schegloff, E. A., Jefferson, G. & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language*, 53(2), 361-382.
- Traverso, V. (2016). *Décrire le français parlé en interaction*. Paris: Ophrys.
- Zorzi, D. (2001). The pedagogic use of spoken corpora: Learning discourse markers in Italian. In G. Aston (éd.), *Learning with corpora* (pp. 85-107). Houston: Athelstan.

* Ces références sont disponibles en ligne



Annexes

Conventions de transcription

Pour la notation des productions verbales, nous avons utilisé les conventions de transcription ICOR, dont une version développée est consultable à l'adresse <http://icar.cnrs.fr/projets/corinte/>:

[]	début et fin de segment en chevauchement	(.)	micro-pause
(2.1)	pause en secondes	xxx	segment inaudible
/ \	intonation montante/ descendante\	exTRA	segment accentué
((rire))	actions à l'écran, phénomènes divers	:	allongement vocalique
< >	délimitation des phénomènes entre (())	par-	truncation
&	continuation du tour de parole	=	enchaînement rapide
^	liaison	.h	aspiration

Pour la multimodalité, nous nous inspirons des conventions développées par Mondada et consultables sur son site personnel (<https://www.lorenzamondada.net/multimodal-transcription>):

Le verbal apparaît en gras pour le distinguer de la description de la multimodalité. Les gestes et les mouvements du corps sont repérés par rapport à une production verbale:

, £, +	indique le début et la fin d'un geste (, £) ou d'un regard (+) décrit à la ligne suivante
---	signale la durée du geste ou du regard
--->	indique que le geste ou le regard continue aux lignes suivantes jusqu'à la prochaine borne (*, £, +)
--->>	indique que le geste ou le regard continue au-delà de la fin de l'extrait
#1	dans le texte, le signe # suivi d'un chiffre indique l'emplacement et le numéro de l'image