



**HAL**  
open science

## La planification des unités prosodiques dans la narration : contrôle intentionnel et contraintes opérationnelles

Monique Vion, Annie Colas

► **To cite this version:**

Monique Vion, Annie Colas. La planification des unités prosodiques dans la narration : contrôle intentionnel et contraintes opérationnelles. Travaux interdisciplinaires du Laboratoire Parole et Langage, 2005, 24, pp.181-196. hal-00134208

**HAL Id: hal-00134208**

**<https://hal.science/hal-00134208>**

Submitted on 1 Mar 2007

**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.

**LA PLANIFICATION DES UNITES PROSODIQUES  
DANS LA NARRATION :  
contrôle intentionnel et contraintes opérationnelles**

**Monique VION et Annie COLAS\***

\*Laboratoire Parole et Langage CNRS– UMR 6057  
Université de Provence  
29 av. R. Schuman  
F-13621 Aix en Provence cedex  
E-mail : [mvion@up.univ-mrs.fr](mailto:mvion@up.univ-mrs.fr)

## RESUME

Des enregistrements de narrations de bandes dessinées sans texte ont été soumis au jugement perceptif d'auditeurs tout venant. Ils ont été invités à les découper en segments prosodiques avec pour critère la variation mélodique et à décider du caractère continuatif ou terminal de chaque segment.

L'hypothèse est que la façon dont un texte oral est divisé en segments de discours par les moyens prosodiques, porte la trace de l'activité cognitive du locuteur et que cette division permet à l'auditeur de structurer sémantiquement à son tour les paroles entendues.

Des enfants francophones de 7 à 11 ans ont raconté les bandes dessinées à un pair d'âge naïf et silencieux. L'établissement de relations temporelles et/ou causales entre les événements était plus ou moins guidé par les dessins (séquences ordonnées vs. séquences arbitraires). La compréhension des liens référentiels était plus ou moins favorisée par le mode de présentation (présentation simultanée vs. consécutive des images). Quatre conditions de production ont été ainsi définies qui diffèrent par les contraintes sur la recherche mnémonique et la résolution inférentielle imposées aux locuteurs.

Les résultats présentés ici concernent l'analyse auditive des productions des enfants de 11 ans. Ils montrent l'importance du mode de présentation des images dans la production de frontières prosodiques terminales.

MOTS CLES : narrations, enfants, unités prosodiques, analyse auditive.

## ABSTRACT

Tape-recorded narrations of word-less picture stories told by native speakers of French were submitted to a perceptual evaluation. Naïve French listeners were invited: 1) to segment them using melodic variation as a criterion; 2) to decide whether each prosodic segment was a « conclusive » or a « continuative » one. The assumption is that the way in which a stretch of speech is divided into segments by means of prosody, carries the trace of speakers' cognitive activity and that this segmentation provides perceptual cues to the listeners to uncover the discourse structure. Children (aged 7 to 11) told their story to a silent same-age peer. The establishment of temporal and/or causal relations between the events was more or less guided by the drawings (ordered vs. arbitrary sequences). The comprehension of the referential links was more or less supported by the way the frames were displayed (simultaneous vs. consecutive display mode). Four conditions of production were thus defined which differed by the constraints on inference-resolving and memory-searching. The present results of the auditory analysis relate to the productions of the 11 year old children. It shows the role of the display mode in the production of conclusive segments.

KEY WORDS : narratives, elementary-schooled children, prosodic units, auditory analyse.

L'idée que la prosodie de la parole spontanée est composée d'unités qui marquent les groupements prosodiques dans le discours fait l'objet d'un consensus notable (Hirst & Di Cristo, 1998). La taille de ces unités caractérisées par une organisation intonative cohérente, est variable. En référence à la composante textuelle de la construction du discours, leur taille peut aller de la simple syllabe à des constituants syntaxiques, des propositions, des phrases ou des productions plus longues (BOTINIS *et al*, 2001).

La question posée à leur égard est celle de savoir ce qui détermine leur construction en ligne par le locuteur lorsqu'il produit son discours. La présente recherche tente une approche psycholinguistique des déterminants de la réalisation de ces unités. Elle cherche à identifier empiriquement certaines des contraintes que l'on suppose s'exercer sur la façon dont le locuteur divise la suite de paroles en segments de discours par les moyens prosodiques.

Lorsqu'il parle, le locuteur verbalise une unité d'information temporairement active dans son esprit. Selon l'hypothèse du caractère incrémental de la production verbale, les unités d'information qui sont mentalement disponibles pour composer le contenu d'un message se présentent comme une séquence cumulée et ordonnée de fragments de ce contenu. Ces fragments de contenu sont verbalisés les uns à la suite des autres, après avoir été traités d'étage en étage (de haut en bas) par les différentes composantes du système de traitement (depuis la conception jusqu'à l'articulation en passant par les composantes grammaticales et phonologiques, LEVELT, 1989 ; BOCK, 1995).

La segmentation prosodique peut avoir des origines diverses. LEVELT (1989) en répertorie cinq relatives à l'intention du locuteur, à la programmation articulatoire, à la syntaxe, à la sémantique et aux contraintes opérationnelles. Deux sont étudiées dans cette recherche.

La première concerne le contrôle intentionnel exercé par le locuteur sur son discours : le locuteur peut se servir des unités prosodiques comme d'un dispositif pragmatique destiné à rendre son discours plus intelligible pour l'auditeur. DI CRISTO (2000) par exemple, défend l'idée que la fonction générale de la prosodie est avant tout une fonction d'assistance à l'encodage et au décodage de la parole qui opère essentiellement au niveau de l'interprétation sémantique et pragmatique des énoncés. Selon ce point de vue, les ruptures formellement introduites dans le discours par la génération des unités prosodiques constituent des frontières. A savoir des marques fonctionnelles qui indiquent l'achèvement (momentané ou définitif) à la fois d'un événement acoustique et de ce que le locuteur tient à dire.

La seconde concerne les contraintes opérationnelles. LEVELT (1989) par exemple envisage que lors du traitement incrémental, une rupture prosodique puisse survenir parce qu'elle permet d'attendre qu'un bloc de matériel soit disponible pour l'encodage phonologique. Le manque d'alimentation peut provenir d'une difficulté de planification du message ou d'une difficulté située au niveau de l'encodage grammatical.

La recherche en cours se fonde sur une double hypothèse : d'une part, la façon dont la suite de paroles est divisée en segments de discours par les moyens prosodiques porte la trace de l'activité cognitive à l'œuvre dans sa planification ; d'autre part, les démarcations prosodiques ainsi réalisées sont exploitées par l'auditeur pour comprendre et interpréter le discours au fur et à mesure de sa réception. Il s'agit donc de montrer que : 1) l'organisation prosodique varie en fonction du contrôle intentionnel exercé par le locuteur (contrôle dépendant de ses compétences linguistiques et communicatives) ainsi qu'en fonction de contraintes opérationnelles (liées à la recherche mnémonique et à la résolution inférentielle) et que 2) cette organisation prosodique permet à l'auditeur de structurer à son tour le discours entendu. Le domaine de la recherche est celui du développement de la narration orale. Le matériel de l'étude est constitué de narrations recueillies sur la base d'un support imagé, auprès d'enfants francophones scolarisés en cycle élémentaire.

Les enregistrements ont été effectués en invitant les enfants (âgés de 7 à 11 ans) à raconter des bandes dessinées sans texte, à un pair d'âge qui ne les connaissait pas. Avec pour l'hypothèse générale que les dispositifs linguistiques employés pour structurer le discours sont la manifestation de contraintes de nature conceptuelle, inhérentes à la gestion de l'information de la part du locuteur (BRONCKART, 1985, CHAFFE, 1986), quatre conditions de production ont été définies. Elles diffèrent par les contraintes cognitives qu'elles imposent aux locuteurs. Il a été fait en sorte que l'établissement de relations temporelles et/ou causales entre les événements représentés sur chaque image soit plus ou moins guidé par ce qui figure dans les dessins (contraintes sur l'activité d'inférence). Il a été fait en sorte par ailleurs que la compréhension des liens référentiels soit plus ou moins favorisée par le mode de présentation des images (contraintes sur l'activité mnémonique). Le croisement de ces deux dimensions (le lien unissant les événements à relater et l'empan des contenus disponibles pour la planification) a permis dans un premier temps de comparer les

productions des enfants du point de vue de leur organisation textuelle (VION & COLAS, 1998, 1999 a et b, 2005).

L'objectif de la présente recherche est d'analyser ces productions du point de vue de leur organisation prosodique, sachant que les variations individuelles inter-locuteurs concernant la configuration d'indices supra-segmentaux utilisés peuvent être très grandes. Le résultat de l'analyse de la composante textuelle (gestion de la co-référence et emploi de connecteurs) constitue le contexte dans lequel l'étude de l'organisation prosodique est entreprise. Il s'agit de montrer que le découpage en unités prosodiques des enregistrements recueillis dépend : 1) du développement des compétences narratives des jeunes locuteurs ; 2) du lien unissant les événements à relater ; et 3) de l'empan des contenus disponibles pour la planification de la production.

L'emploi d'une tâche de jugement perceptif pour identifier l'organisation prosodique est motivé par l'hypothèse qu'un auditeur peut mobiliser de façon réfléchie et consciente sa capacité à se fonder spontanément sur les indices prosodiques présents dans le discours pour recouvrer sa structuration. Les études empiriques sur ce point sont encore peu nombreuses et prennent peu en compte les variations inter-locuteurs (voir les bilans établis par CUTLER *et al.*, et FON, 2002). Mais ces études montrent d'ores et déjà que l'information mélodique (les variations de FO et de la gamme mélodique, les accents de syntagmes et les tons de frontière) est exploitée pour traiter les types spécifiques de relations existant entre l'énoncé en cours et ceux qui le précèdent et le suivent (HIRSCHBERG & PIERREHUMBERT, 1986 ; PIERREHUMBERT & HIRSCHBERG, 1990 ; GROSZ & HIRSCHBERG, 1992 ; Grosz *et al.*, 1994 ; SWERTS & GELUYKENS, 1994 ; AURAN, *et al.*, 2001, PORTES, 2002).

L'analyse auditive de l'organisation prosodique des narrations a été confiée à des juges non experts en linguistique. Ces juges novices ont été invités : 1) à découper chaque enregistrement en segments en se fondant sur la mélodie de la parole ; 2) à décider de la finalité ou de la continuité de chacun ; et 3) à estimer la force de séparation des démarcations ainsi identifiées. La démarche se fonde sur l'idée avancée par ROTONDO (1984) et reprise par SWERTS (1997) selon laquelle les segmentations subjectives fournies par un groupe d'auditeurs sont exploitables pour dégager la hiérarchie des frontières de discours. La proportion de sujets s'accordant sur une rupture donnée fournissant une indication sur la force de séparation de cette frontière dans le discours. Elle se fonde également sur le résultat de l'étude de AURAN *et al.* (2001) qui montre que les auditeurs sont capables

d'opérer une distinction catégorielle entre « continuité » et « conclusivité » dans le discours oral spontané.

Eu égard à la variété d'organisation des textes résultant de l'interaction des contraintes mnémoniques et inférentielles manipulées, on peut s'attendre à observer des différences, mais aussi des régularités dans l'organisation prosodique des productions. Régularités liées pour une part aux contraintes opérationnelles manipulées et pour une autre part à l'évolution des compétences communicatives des enfants (contrôle intentionnel). Les attentes qui seront détaillées ultérieurement en fonction de l'âge et des conditions de production tiennent compte du poids exercé par ce double jeu de contraintes.

## METHODE

### 1. Matériel

Les enregistrements soumis au jugement sont issus du recueil effectué par les auteurs sur la base de bandes dessinées composées de 8 vignettes de format 8 x 8 cm (VION & COLAS, 2000). Chaque bande dessinée se présentait sous deux formes. Dans l'une, les images étaient disposées sur une seule page (présentation *simultanée* – S- des images). Dans l'autre, la bande dessinée se présentait sous la forme d'un carnet à raison d'une image par page (présentation *consécutive* – C- des images). Les enregistrements ont été effectués à l'école dans une pièce calme, mais non isolée des bruits environnants. La passation était individuelle. Trois personnes étaient présentes : le locuteur, l'expérimentateur et le destinataire de la narration. Le destinataire était un pair d'âge de même niveau scolaire que le locuteur. Il n'était auditeur qu'une seule fois au cours du recueil. Il était suffisamment éloigné du locuteur pour ne pas voir les images, de sorte que ce dernier ne puisse pas fonder sa narration sur une connaissance partagée. Chaque locuteur était confronté à un seul mode de présentation. En présentation simultanée, le locuteur était invité à prendre connaissance de la bande dessinée afin de la raconter immédiatement après. En présentation consécutive, le locuteur était invité à tourner les pages du carnet une à une et à raconter au fur et à mesure de la découverte des images. Dans les deux cas, la consigne donnée au destinataire était la même. Il devait être attentif : il devait écouter les histoires pour bien les comprendre, sans parler.

Les enregistrements retenus pour la présente recherche ont pour support deux bandes dessinées dans lesquelles un personnage est engagé dans diverses activités. Ces bandes dessinées s'opposent selon le caractère plus ou moins explicite par l'image de l'enchaînement des événements figurés sur la séquence de vignettes (Figures 1). Dans l'une (Figure 1a), les activités représentées sont relativement indépendantes les unes des autres et hautement sujettes à inférence. L'image de la femme qui s'habille (ou se déshabille) avant de se maquiller (ou de se démaquiller) aurait pu se trouver après l'image de la femme au miroir ou dans n'importe quelle autre position dans la séquence (séquence *arbitraire* - A). Dans l'autre (Figure 1b), un lien séquentiel minimal entre les images est fourni. S'équiper pour la pêche, aller au bord de l'eau, y plonger une ligne sont des activités qui s'enchaînent ordinairement avant de faire éventuellement une capture (séquence *ordonnée* - O). L'enchaînement ordinaire des événements est dévié par le surgissement d'un obstacle qui est un événement externe sur lequel le personnage principal n'a pas de contrôle. Chaque locuteur était confronté à un seul type de séquence.

/insérer ici les figures 1/

En résumé, les enregistrements sélectionnés pour constituer le matériel de la présente recherche ont été effectués dans l'une des quatre conditions de production issues du croisement des dimensions « mode de présentation » et « type de séquence » (soit : AC, AS, OC et OS). Ils proviennent de quatre-vingt-quatre locuteurs différents (garçons et filles). Soit 28 enfants de 7 ans, 28 enfants de 9 ans et 28 enfants de 11 ans. Les productions, d'abord enregistrées sur bande magnétique, ont été ensuite numérisées et stockées au moyen du logiciel SoundEdit 16, à raison d'un fichier-son par production. Les 84 productions représentent au total vingt et une minutes d'enregistrement.

## **2. Tâche de segmentation pondérée**

Le découpage des productions en segments prosodiques repose sur une analyse perceptive des enregistrements semblable à celle qu'ont pratiquée PORTES (2000, 2002) et AURAN et al., (2001). La tâche de segmentation pondérée mise au point pour l'expérience est plus directement inspirée de PORTES (2000, 2002). Les juges effectuent ici la segmentation puis la pondération « à la volée ». Chaque juge dispose d'un système binaire de réponse qui se compose d'une barre oblique (/) pour délimiter un segment qui laisse le message en suspens

et une double barre oblique (//) pour délimiter un segment qui induit un effet de clôture. Pour pondérer les juges disposent d'une échelle en quatre points, où « 1 » indique que la rupture identifiée a une force de séparation faible et « 4 » une séparation très forte. L'échelle est inspirée du système ToBI (« Tones and Break Indices » : SILVERMAN, *et al.* 1992), standard en vue d'étiqueter les traits prosodiques de corpus informatisés en anglais sur un échelle en cinq points (de 0 à 4, où 0 signale une absence de frontière).

### **3. Organisation du recueil**

Compte tenu de la durée totale des enregistrements et du coût attentionnel de la tâche à effectuer, un même juge ne pouvait pas traiter la totalité du matériel. L'ensemble des enregistrements a donc été divisé en 14 blocs comportant chacun six enregistrements. Dans un bloc, les six productions ont été choisies de telle sorte qu'un juge se trouve confronté à des enregistrements concernant les trois âges (7 ans, 9 ans et 11 ans), les deux types séquences (A et O) et les deux modes de présentation (C et S). Ceci conduit à deux compositions complémentaires des blocs. Sept blocs obéissent à la composition : [9OC - 7OS - 11OS - 9 AS - 11AC - 7AC] et sept blocs à la composition : [9OS - 7OC - 11OC - 9AC - 11AS - 7 AS]. Les juges entendent les six enregistrements d'un bloc soit dans l'ordre indiqué ci-dessus, soit dans l'ordre inverse (chaque production était donc proposée au sein de deux listes). La segmentation pondérée de la totalité des enregistrements a nécessité la constitution de 28 listes où six enregistrements se succèdent. Quatre juges ont participé à la segmentation pondérée des enregistrements d'une liste.

### **4. Participants**

Les juges sont des auditeurs non experts en linguistique. Il s'agit de jeunes adultes, étudiants en lettres à l'Université de Provence, ayant le français pour langue maternelle. Au total la segmentation pondérée des enregistrements a nécessité la participation de cent douze juges (4 juges x 28 listes).

### **5. Procédure**

La tâche est effectuée dans un local insonorisé. Les quatre juges affectés à une liste sont munis d'écouteurs et effectuent individuellement, mais simultanément la tâche qui est pilotée par ordinateur. Le programme, écrit grâce au logiciel PsyScope (COHEN *et al.*, 1993), assure la gestion et la présentation des fichiers-son.

L'expérimentateur commence d'abord par indiquer aux juges l'objet de la recherche. Puis, pour les familiariser avec la tâche, il leur propose d'effectuer un essai d'entraînement décomposé en quatre phases. Une première phase consiste à écouter l'enregistrement pour en prendre connaissance. A ce stade, aucun appui visuel n'est fourni : les juges ne disposent ni de la transcription correspondante, ni de la bande dessinée. Les juges reçoivent ensuite une transcription orthographique de l'enregistrement qu'ils viennent d'entendre. La transcription n'est pas ponctuée et ne comporte pas de majuscules. Les juges sont invités à une seconde écoute au cours de laquelle ils noteront au fur et à mesure sur la transcription le découpage perçu (phase 2). Lors d'une troisième écoute (phase 3), ils sont invités à indiquer sur la transcription la force des ruptures notées dans la phase précédente. Dans la quatrième phase, une dernière écoute de l'enregistrement permet de vérifier les jugements et éventuellement de les modifier. Ces mêmes phases sont reproduites ensuite pour traiter les 6 productions dévolues à chaque juge dans l'expérience proprement dite.

L'expérimentateur dispense des temps de repos entre chaque phase de la tâche en fonction de la demande du groupe. Le temps de réalisation de la tâche n'excède pas 30 minutes.

A l'issue de l'expérience, les juges sont invités à indiquer par écrit quels sont les critères qu'ils ont utilisés pour décider de la poursuite ou de l'achèvement d'un énoncé au-delà de la rupture identifiée et pour indiquer sa force de séparation.

## **6. Attentes**

Les données analysées dans le cadre du présent article concernent le résultat du jugement des enregistrements des enfants de onze ans. Du point de vue du contrôle intentionnel, ces locuteurs se sont révélés les plus à même de s'engager à raconter les histoires proposées en BD plutôt qu'à décrire une à une les images qui les composent et donc d'en rapporter le contenu en planifiant leur production dans un cadre unifié. Sous l'hypothèse que les auditeurs-juges recouvrent la segmentation du discours sur la base de l'organisation prosodique réalisée, les attentes concernent l'effet des contraintes inférentielles et mnésiques (opérialisées respectivement par les variables « mode de présentation » et « type de séquence ») sur l'organisation prosodique des productions des enfants.

De nombreux auteurs s'accordent sur l'idée que le « phrasé » prosodique est corrélé avec les unités d'information. Les bandes dessinées comportant huit images et le contenu de chaque image pouvant être énoncé en une seule proposition, on peut s'attendre à ce que les locuteurs réalisent au minimum huit unités prosodiques par enregistrement. Mais le nombre

d'unités réalisées ainsi que le caractère (continuatif vs. terminal) qui leur est associé devrait varier en fonction du jeu combiné des variables manipulées.

La présentation *simultanée* des images ainsi que la séquence *ordonnée* sont en effet moins coûteuses du point de vue de la mobilisation des ressources cognitives du locuteur que la présentation *consécutive* et la séquence *arbitraire*. L'étude des productions dont sont extraites celles du présent corpus a montré que la présentation simultanée des images ainsi que le caractère ordonné favorisaient davantage l'engagement des locuteurs dans un processus narratif que les deux autres modalités.

Sur la base d'une présentation simultanée des images et/ou de la séquence ordonnée d'événements, le locuteur peut ou bien agencer le contenu de plusieurs images et le concentrer en une seule proposition (par exemple, dire pour les images 2 à 5 de la séquence ordonnée en présentation simultanée : « le garçon s'installe pour pêcher ») ou bien au contraire rapporter ce même contenu en plusieurs propositions (« le garçon s'approche de la rivière, s'assoit avec son épuisette à côté de lui, trempe sa canne à pêche dans l'eau et attend »). Il peut éventuellement ajouter des informations sur les états internes ou les motivations du personnage (« il est content »). Dans ces circonstances, le nombre de ruptures prosodiques ainsi que le caractère (continuatif vs. terminal) qui leur est reconnu devrait se révéler plus variable d'un locuteur à l'autre que celui observé pour les locuteurs qui découvrent les images au fur et à mesure ou qui sont confrontés à la séquence arbitraire, modalités qui induisent plus fortement le traitement isolé du contenu de chaque image. On s'attend donc à observer un effet d'interaction des variables manipulées sur la segmentation perçue par les auditeurs.

## RESULTATS

Les données analysées concernent le jugement de 28 productions. Soit sept narrations par condition de production soumises chacune à l'analyse auditive de huit juges. Pour chaque narration, la confrontation des découpages perçus par chacun de ses auditeurs a permis d'établir un découpage unique qui rassemble les informations concernant les ruptures, leur caractère (continuatif vs terminal) et la pondération associée à chacune d'elles.

### 1. Découpage retenu pour l'analyse

Pour dégager les frontières de discours empiriquement perçues, il a été procédé au calcul de la proportion de juges qui s'accordent sur la présence d'une rupture donnée. Rappelons que cette démarche se fonde sur l'hypothèse selon laquelle on peut dériver la structure hiérarchique du discours de la segmentation effectuée par plusieurs auditeurs. Les désaccords entre les juges sur les ruptures les plus fortes étant d'autant moins importants que ces ruptures présentent des transitions claires dans le flux d'information. L'obtention de ce découpage unique, repose sur les opérations suivantes.

La première opération a consisté à éliminer au hasard le protocole de l'un des huit juges, de façon à disposer d'un nombre de jugement maximal impair qui permette de retenir sans ambiguïté les ruptures identifiées à la majorité (c'est-à-dire : indiquées par au moins quatre juges sur sept). Ainsi, pour qu'un fait soit retenu pour l'analyse, un seuil minimal d'accord entre les juges de 57 % doit être atteint.

Pour chaque enregistrement donc, la présence d'une rupture n'a été reportée manuellement sur la transcription servant de base aux analyses statistiques ultérieures que si elle avait été indiquée par au moins 57% des juges. Ce report est assorti de l'indication du nombre de juges l'ayant identifiée. Ce nombre peut varier entre 4 et 7.

Dans ce report, chaque rupture porte la mention de son caractère continuatif (/ ) vs. Terminal (//). Il correspond à celui indiqué par le plus grand nombre de juges ayant noté la présence de la rupture. Ce nombre qui peut varier entre 2 et 4 est également noté sur la transcription. La pondération associée à la rupture ainsi caractérisée correspond à la moyenne des poids exprimés par cette majorité de juges.

## **2. Codage des transcriptions**

La transcription correspondant à chaque enregistrement a été ensuite codée en vue du tri automatique des données. Un code à cinq places a été associé à chaque rupture. Une place permet de noter le nombre de juges ayant indiqué la présence de la rupture. Une autre place concerne son type (noté c pour / et t pour //). Une autre permet d'indiquer le nombre de juges l'ayant ainsi noté. Une autre encore permet de noter la force de séparation attribuée à cette rupture. Une autre enfin permet de noter le type (C ou T) qui a été retenu pour cette même rupture par les auteurs de la recherche, antérieurement au recueil des jugements. Cette dernière information a permis de départager le choix des juges en cas d'équirépartition des jugements concernant la caractérisation de la rupture. Dans ce cas, la force moyenne de

la rupture est calculée en se fondant sur la pondération des juges qui partageaient l'avis des auteurs.

Les transcriptions ainsi codées ont été ensuite éditées selon les conventions du système CHILDES (Child Data Exchange System, MACWHINNEY, 1991) qui permet d'effectuer automatiquement les recherches d'occurrences.

### 3. Analyses

Diverses analyses ont été conduites sur la base de variables dépendantes dérivées des informations relatives aux ruptures perçues et au type de chacune. Pour étudier l'effet des contraintes inférentielles et mnésiques sur ces variables dépendantes, des analyses de la variance ont été effectuées avec les facteurs « Mode de Présentation » et « Type de Séquence », comme facteurs intra-sujet. Ce qui peut s'exprimer par la formule : Enregistrements  $_7$  < Mode de Présentation  $_2$  \* Type de Séquence  $_2$ >. Seuls les effets significatifs à un seuil inférieur ou égal à .05 sont pris en compte pour les commentaires.

En préalable à ces études, la durée des productions a été examinée. L'analyse de la variance conduite avec la durée comme variable dépendante montre que celle-ci ne diffère pas significativement en fonction des deux variables manipulées ni de leur interaction (leur durée moyenne est de 46 secondes).

De la même manière, le nombre total de ruptures par enregistrement a été examiné. Ce nombre, retenu en appliquant le critère d'un avis majoritaire, est en moyenne de 14,07. Il est à noter que dans un enregistrement, en moyenne 61,43 % de ces ruptures sont consensuelles (unanimité des juges).

L'accord entre les juges et les auteurs de la recherche en ce qui concerne l'appréciation du caractère terminal ou continuatif de chaque rupture a été aussi examiné. Il est pour l'ensemble du corpus égal à 85,79 %.

Les analyses conduites avec pour variables dépendantes le nombre de ruptures perçues par enregistrement, le nombre de frontières continuatives et le nombre de frontières terminales montrent une absence de variation significative en fonction des variables manipulées. L'absence de résultats significatifs à ce niveau d'appréhension des données a conduit à restreindre, puis à affiner les données soumises à l'analyse.

Dans un premier temps, les ruptures identifiées dans la partie de l'enregistrement correspondant à la description de la première image ont été exclues. Ceci a été fait sous l'hypothèse que cette partie de la production, qui correspond à l'exposé du cadre et à la présentation des personnages, a pu à elle seule introduire une grande variabilité dans les données à traiter. Dans la partie des enregistrements retenue (qui concerne donc la production se rapportant à la séquence des images 2 à 8 où un seul personnage reste en scène), le nombre de ruptures majoritairement identifiées par les juges est de 13 en moyenne. Il est à noter que 62 % de ces ruptures sont consensuelles. L'analyse conduite sur ces données montre là aussi que le nombre de ruptures ne varie pas de façon significative en fonction des variables manipulées ni de leur interaction.

L'analyse a alors été poursuivie en élagant ces données de la façon qui suit. L'opération a consisté à distinguer des autres, parmi les ruptures identifiées à la majorité des juges, celles qui correspondaient à l'emploi d'expressions phatiques (« heu » ; « alors heu » ; « donc heu ») ou à la mise en mots (« i prend/ i prend la/i prend le filet » ; « alors/ c'est / le mari/ de /enfin ... » ) ou plus rarement de la mise en phrase (« iii il prend un filet puis après va / il le /il essaye de pêcher »). Ceci a été fait sous l'hypothèse que ces ruptures témoignent d'activités qui n'altèrent pas le projet communicatif général du locuteur. Cette dichotomie a été pratiquée indépendamment par chacun des auteurs. Les rares désaccords qui ont surgi lors de la confrontation du résultat de l'opération ont été résolus par la discussion.

Cette dichotomie a permis de faire apparaître un premier ensemble de faits. En premier lieu, les ruptures de cette catégorie sont globalement peu fréquentes (elles représentent 7,05 % des ruptures identifiées dans la production se rapportant à la séquence des images 2 à 8). En second lieu, ces ruptures ne sont pas également représentées dans les quatre conditions de production. Il y a plus d'enregistrements porteurs de ruptures liées à l'emploi d'expressions phatiques et/ou à la mise en mots en présentation simultanée (9 enregistrements sur 14) qu'en présentation consécutive (3 enregistrements sur 14). Bien que ne pouvant être testée au plan statistique, cette observation va dans le sens d'une incidence du mode de présentation des images sur la présence de cette catégorie de rupture qui peut être vue comme relevant d'un comportement de temporisation lié à une tentative de planification à long terme.

Les analyses suivantes ont porté sur la catégorie complémentaire. A savoir : les ruptures non liées à la fonction phatique et/ou aux choix lexico-syntaxiques. L'analyse conduite avec comme variable dépendante le nombre de ruptures identifiées par la majorité des juges montre une absence de variation significative en fonction des variables manipulées. En revanche, l'analyse conduite avec les seules ruptures consensuelles fait apparaître une tendance, très proche de la valeur seuil fixée pour l'étude [interaction Type de Séquence x Mode de Présentation :  $F(1, 24) = 4,087$  ;  $p < .0545$ ]. En présentation simultanée, les juges ont identifié moins de ruptures dans la production des locuteurs confrontés à la séquence ordonnée que dans celle des locuteurs confrontés à la séquence arbitraire (5,9 vs. 7,3). L'inverse s'est produit en présentation consécutive où les juges ont perçu davantage de ruptures dans la production des locuteurs confrontés à la séquence ordonnée que dans celle des locuteurs confrontés à la séquence arbitraire (8 vs. 6,7).

L'analyse a été ensuite poursuivie sur cette même catégorie de rupture en examinant le caractère (continuatif vs. terminal) qui leur a été reconnu. Elle a consisté à distinguer les ruptures que les juges avaient majoritairement identifiées comme des frontières terminales de celles qu'ils avaient majoritairement identifiées comme des frontières continuatives.

L'analyse conduite avec le nombre de frontières continuatives identifiées par la majorité des juges comme variable dépendante montre que ce nombre ne varie pas significativement en fonction des variables manipulées ni de leur interaction. En revanche, l'analyse conduite avec le nombre de frontières terminales identifiées par la majorité des juges comme variable dépendante indique un effet significatif du mode de présentation [  $F(1,24) = 4,292$  ;  $p < .0492$ ]. Les frontières terminales ont paru plus nombreuses aux juges dans les enregistrements recueillis en présentation consécutive (3,07 unités en moyenne) que dans les enregistrements recueillis en présentation simultanée (2,07 unités en moyenne). Ce résultat conforte l'idée que la présentation consécutive favorise la production d'unités informationnelles associées à des schémas prosodiques induisant un effet de clôture.

Une nouvelle analyse (avec le nombre de frontières terminales identifiées par la majorité des juges comme variable dépendante) a été conduite en ne conservant que les productions dont la composante textuelle attestait que les locuteurs étaient véritablement engagés dans un processus narratif (caractérisé par la prise en compte du caractère progressivement accru d'ancienneté du personnage et par une diversification du marquage explicite des liens). Ceci a conduit à exclure de l'analyse deux enregistrements caractérisés par le fait que le personnage

était toujours désigné par la même expression référentielle (« la dame »). Cette dernière analyse indique un effet significatif accru du mode de présentation sur l'appréciation donnée par les juges de la présence de frontières terminales [  $F(1,22) = 7,315$  ;  $p < .0129$  ].

## DISCUSSION

Rappelons que les résultats présentés ci-dessus ne concernent qu'une partie des données que le projet de recherche envisage de traiter. Il est donc trop tôt encore pour se pencher sur la question du contrôle intentionnel opérationnalisé dans la recherche par la variable « âge ». En revanche, les résultats issus de l'analyse des productions des locuteurs les plus âgés (onze ans) montrent l'impact de la dimension mnésique sur la réalisation et la reconnaissance des frontières terminales. Les auditeurs ont en effet identifié davantage de frontières terminales dans la production du groupe de locuteurs ayant raconté au fur et à mesure de la prise de connaissance des images que dans la production du groupe de locuteurs ayant pris connaissance de la totalité de la bande dessinée antérieurement à la narration. Ce qui souligne le fait qu'en présentation consécutive les locuteurs ont marqué les informations plutôt comme des entités clôturées que comme des entités appelant une suite. Les résultats suggèrent également que cette dimension entre en interaction avec la dimension inférentielle. Ces deux dimensions ont été manipulées sous l'hypothèse qu'une présentation simultanée des images (qui favorise une représentation d'ensemble des événements à relater) et/ou un enchaînement ordonné d'actions (qui facilite l'appréhension d'un sujet thématique) sollicitent moins les capacités mnémoniques et/ou inférentielles du narrateur qu'une présentation au fur et à mesure ou un enchaînement arbitraire des images. Il semble que ces composantes cognitives de la situation de production, qui favorisent ou ne favorisent pas la création d'une représentation conceptuelle d'ensemble préalable à la mise en forme narrative, puissent avoir au moins une incidence quantitative sur l'organisation prosodique réalisée.

Ces résultats partiels sont plutôt encourageants pour la poursuite et la systématisation d'une approche expérimentale des déterminants de la segmentation prosodique. Mais ils permettent aussi de prendre la mesure du prix à payer pour les obtenir. C'est en effet au prix de transformations, de restrictions et d'affinements des données originales. Toutes ces

opérations demandent de prendre des décisions théoriquement justifiées tant au plan linguistique que psychologique.

## REFERENCES

AURAN, C., PORTES, C., RAMI, E., & RIGAUD, N. (2001). *La distinction entre frontières conclusives et continuatives est-elle pertinente dans le discours spontané ?* Paper presented at the "Journées Prosodie", Grenoble.

BOCK, K. (1995). Sentence production: from mind to mouth. In P. D. J.L. Miller et Eimas (Ed.), *Speech Language and Communication*. London: Academic Press, p. 181-216..

BOTINIS, A., GRANSTRÖM, B., & MÖBIUS, B. (2001). Developments and paradigms in intonation research. *Speech Communication*, 33, p. 263-296.

BRONCKART, J. P. (1985). Pour un modèle de production du discours. In J. P. Bronckart (Ed.), *Le fonctionnement des discours*. Neuchâtel: Delachaux et Niestlé, p.3-58.

CHAFE, W. (1986). Cognitive constraints on information flow. In R. Tomlin (Ed.), *Coherence and grounding in discourse. Typological studies in language* (Vol. 11). Amsterdam: John Benjamins Publishing Compagny, p. 21-51.

COHEN, J., MAC WHINNEY, B., FLATT, M., & PROVOST, J. (1993). An interactive graphical system for designing and controlling experiments in the psychological laboratory using MacIntosh computers. *Behavior Methods, Research, Instruments and Computer*, 25, p. 257-271.

CUTLER, A., DAHAN, D., & DONSELAAR, W. (1997). Prosody in the comprehension of spoken language: A litterature review. *Language and Speech*, 4, 2, p. 141,201.

DI CRISTO, A. (2000). *Interpréter la prosodie*. Paper presented at the "XXIIIèmes Journées d'Étude sur la Parole", Aussois, p. 241-244.

FON, Y-J. J. (2002). *A cross-linguistic study on syntactic and discourse boundary cues in spontaneous speech*. The Ohio State University.

GROSZ, B., & HIRSCHBERG, J. (1992). *Some intonational characteristics of discourse structure*. Paper presented at the *Proceedings ICSLP* (Banff), p. 429-432.

GROSZ, B., HIRSCHBERG, J. & NAKATANI, G. H. (1994). A study of intonation and discourse structure in directions. Paper presented at the Working papers of the *Workshop on the integration of Natural Language and Speech Processing*.

HIRSCHBERG, J., & PIERREHUMBERT, J. (1986). The intonational structuring of discourse. Paper presented at the Proceedings of the *24th Annual Meeting of the Association for Computational Linguistics*, p. 136-144.

HIRST, D., & DI CRISTO, A. (1998). A survey of intonation systems. In D. Hirst & A. DiCristo (Eds.), *Intonation systems: A survey of twenty languages*. Cambridge: Cambridge University Press , p. 1-44.

LEVELT, J. M. (1989). *Speaking. From intention to articulation*. London: MIT Press.

MACWHINNEY, B. (1991). *The CHILDES project. Tool for analyzing talk*. Hillsdale, N J: Lawrence Erlbaum.

PIERREHUMBERT, J., & HIRSCHBERG, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. R. Cohen , J. Morgan & M. E. Pollack (Eds.), *Intentions in communication* .Cambridge, MA: MIT Press, p. 271-323.

PORTES, C. (2000). *Approche du rôle de la prosodie dans la structuration du discours oral en Français*. Note de Recherche non publiée, Université de Provence, Aix-en-Provence.

PORTES, C. (2002). Approche instrumentale et cognitive de la prosodie du discours en français. *Travaux Interdisciplinaires Parole et Langage*, 21, p.101-119.

ROTONDO, J. A. (1984). Clustering analyses of subjective partitions of text. *Discourse Processes*, 7, p. 69-88.

SILVERMAN, K., BECKMAN, M., PITRELLI, J., OSTENDORF, M., WIGHTMAN, C. M., PRICE, P., PIERREHUMBERT, J., & HIRSCHBERG, J. (1992). ToBI : A standard for labeling English prosody. Paper presented at the *Proceedings ICSLP*, 867-870.

SWERTS, M. (1997). Prosodic features at discourse boundaries of different strength. *The Journal of the Acoustical Society of America*, 101, 1, p. 514-521.

SWERTS, M., & GELUYKENS, R. (1994). Prosody as a marker of information flow in spoken discourse. *Language and Speech*, 37, 1, p. 21-43.

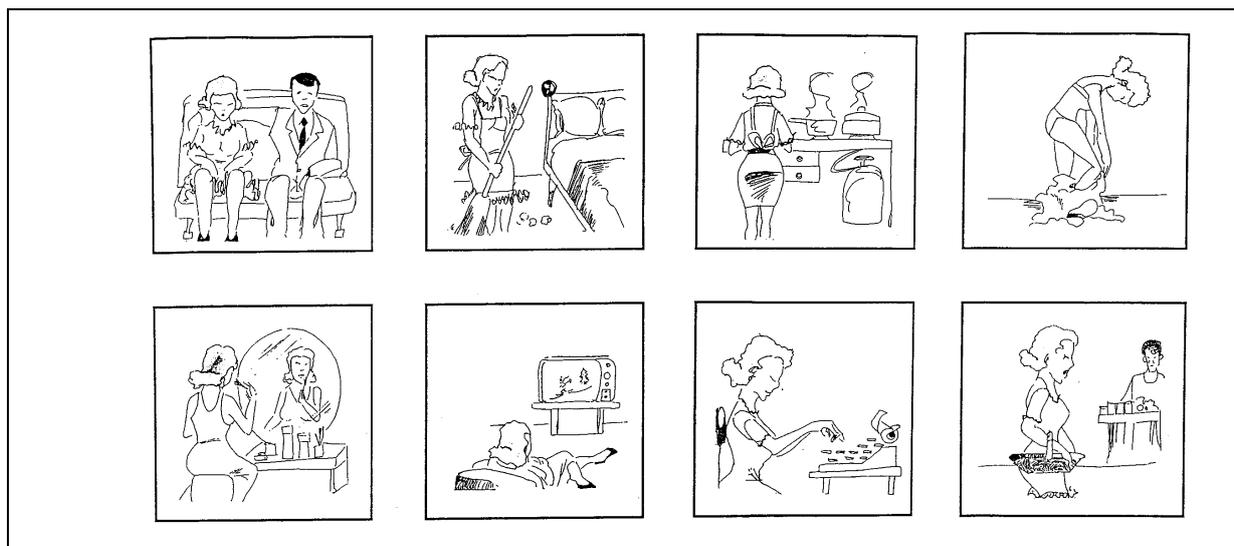
VION, M., & COLAS, A. (1998). L'introduction des référents dans le discours en français : contraintes cognitives et développement des compétences narratives. *L'Année Psychologique*, 98, p. 37-59.

VION, M., & COLAS, A. (1999a). Expressing coreference in French: Cognitive constraints and development of narrative skills. *Journal of Psycholinguistic Research*, 28, 3, p. 261-291.

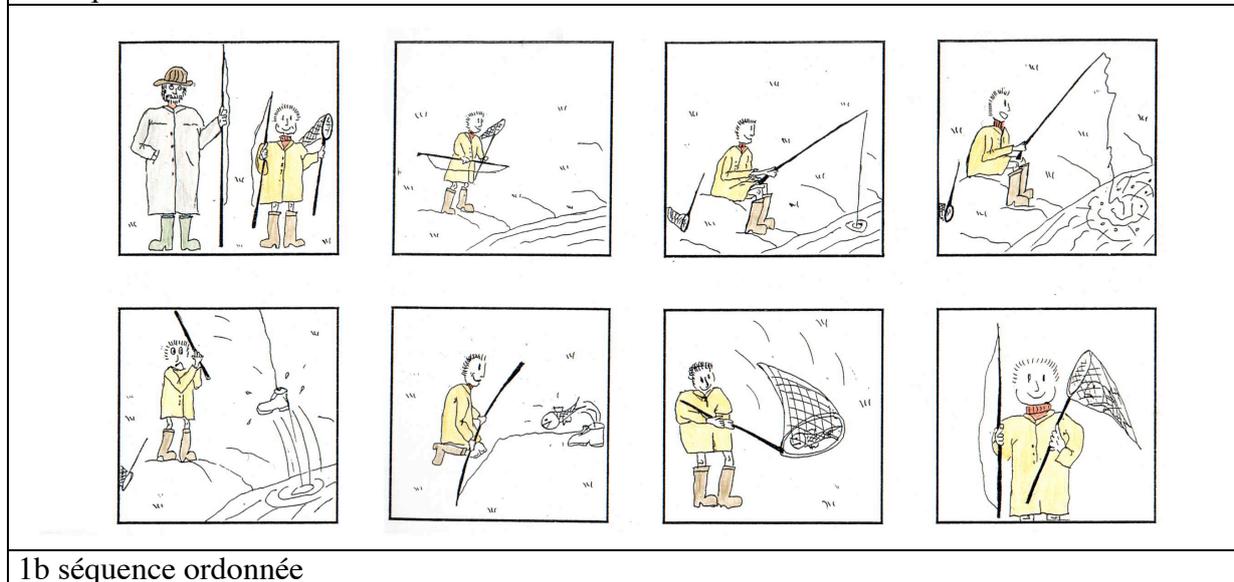
VION, M., & COLAS, A. (1999b). Maintaining and reintroducing referents in French: Cognitive constraints and development of narrative skills. *Journal of Experimental Child Psychology*, 72, p. 32-50.

VION, M., & COLAS, A. (2000). Mode de recueil et outil d'analyse d'un corpus de parole spontanée étudié d'un point de vue psycholinguistique. *Travaux Interdisciplinaires Parole et Langage*, 11, p.155-167.

VION, M., & COLAS, A. (2005). Using connectives in oral French narratives: Cognitive constraints and development of narrative skills. *First Language*, 25 (1), 39-66 .



1a séquence arbitraire



1b séquence ordonnée

Fig. 1 : matériel.