

## Analyse multimodale des particules d'extension « et tout ça, etc. » en français

Gaëlle Ferré, Université de Nantes – LLING

Les *particules d'extension* (aussi appelées *marqueurs finaux de liste* dans Lacheret, 2004) de l'oral soulèvent des questions intéressantes car elles ont été très peu décrites sur le plan de la prosodie, un peu plus sur le plan du discours (voir l'ouvrage de M. Overstreet, 2000, qui décrit le fonctionnement des *adjunctive/disjunctive general extenders* de l'anglais oral), et les études qui portent sur la mise en relation des unités de la prosodie avec celles du discours ou de la syntaxe ne disposent pas de description prosodique détaillée de ce type de particules dont elles ne savent que faire. Il semble de plus que les chercheurs qui décrivent les particules d'extension sur le plan du fonctionnement discursif et syntaxique uniquement aboutissent à des classifications contradictoires et qui font toujours débat à l'heure actuelle. Deux questions se posent concernant ces particules auxquelles nous tenterons d'apporter des réponses : (a) les particules d'extension doivent-elles compter comme marqueurs discursifs/pragmatiques (MD)? Et (b) quelle est leur fonction ?

**CORPUS :** Nous proposons d'aborder ces deux questions en testant la validité de certaines hypothèses déjà émises en analyse de discours lorsque l'on intègre des critères acoustiques et des critères visuels dans une perspective multimodale. Les tests sont effectués sur 103 occurrences des particules « (et) tout (ça) » et « et cetera » sur 3 heures d'enregistrement vidéo tirées du corpus CID, 6 locuteurs (Bertrand et al., 2008). Pour comparaison, le nombre d'occurrences de ces particules dans le CID est légèrement plus élevé que dans d'autres corpus de français parlé (corpus Beeching, CLAPI et Orléans).

**MARQUEURS DISCURSIFS ?** Aucun consensus n'a encore été atteint concernant la classification ou non des particules d'extension comme MD. Les études qui portent sur la description des MD (Schiffrin, 1987 ; Fraser, 1999) les excluent car elles ne répondent pas à l'ensemble des critères : par exemple, elles n'apparaissent jamais en début d'énoncé et leur signification n'est pas toujours éloignée de leur sens propre<sup>9</sup>. Par contre, elles n'apparaissent jamais seules comme sujet ou objet d'un groupe verbal et présentent un grand nombre de contours prosodiques. D'autres études classent des ponctuels tels que « quoi » comme MD, alors que ceux-ci ne répondent pas plus aux critères émis pour la classification des MD (d'ailleurs, lorsqu'ils sont décrits sur le plan prosodique, on leur assigne un statut particulier en les considérant comme extra-métriques). Si l'on suit ces études, on peut être amené à considérer les particules d'extension comme ayant un fonctionnement proche des ponctuels et donc à les classer comme MD. Overstreet (2000) adopte une position intermédiaire car elle considère que les particules d'extension peuvent fonctionner comme MD, mais pas systématiquement. Sa proposition consiste à ne pas opérer de classification générale *a priori* mais à voir le statut de chaque particule en contexte. C'est la position que nous avons adoptée dans cette étude, car certaines d'entre elles nous semblent clairement jouer un rôle méta discursif de ponctuel mais pas toutes. Afin de les distinguer, nous avons adopté un double critère : les particules ont été classées comme MD lorsqu'elles étaient suivies d'une pause (attendue après un ponctuel) et que leur portée sémantique était large. Toutes les autres

---

<sup>9</sup> Ainsi, par exemple, le Petit Robert décrit « et cetera » comme une locution signifiant « et les autres choses » ou « et le reste », et il se trouve que c'est la signification de la particule dans un certain nombre de contextes, même si elle peut avoir d'autres valeurs, comme nous le verrons plus loin.

particules d'extension ont été étiquetées comme locutions. Nous avons ensuite annoté les caractéristiques prosodiques, gestuelles et syntaxiques suivantes pour les particules<sup>10</sup> :

CRITERES PROSODIQUES	CRITERES GESTUELS	CRITERES SYNTAXIQUES
<ul style="list-style-type: none"> <li>• Accentuation</li> <li>• Présence d'une frontière à Gauche et a Droite</li> <li>• Allongement perçu<sup>11</sup> (contexte immédiat et particule)</li> <li>• Plage intonative sur la particule (Haute, Moyenne, Basse)</li> <li>• Mouvement mélodique (montant, plat, descendant)</li> </ul>	<ul style="list-style-type: none"> <li>• Mouvements de tête</li> <li>• Gestes manuels en adoptant la typologie de McNeill (1992)</li> </ul>	<ul style="list-style-type: none"> <li>• Portée syntaxique des particules</li> </ul>

Les résultats montrent que sur le plan syntaxique, il n'y a pas de différence entre les MD et les locutions, dont la portée est majoritairement un VP ou un NP, mais il apparaît que les particules « et cetera » et « tout ça » sont beaucoup plus susceptibles de jouer un rôle de MD que « et tout ça » et « et tout ». Sur le plan prosodique, il apparaît que les MD sont régulièrement désaccentués (ne portent pas l'accent nucléaire, cf. Selkirk, 2003) et ne sont pas allongés. En revanche, peu de MD sont réduits phonétiquement ; les locutions ne le sont jamais. Les MD ne sont le plus souvent pas séparés du syntagme intonatif (Intonational Phrase, Selkirk, 2003) précèdent par une frontière de groupe, même si les marqueurs, tout comme les locutions, montrent un large éventail de contours intonatifs. Ces observations nous permettront de proposer une analyse des MD et des locutions dans le cadre de la théorie du phrasé prosodique de Selkirk (2003).

Sur le plan de la gestualité, les MD sont accompagnés de gestes manuels (plus particulièrement des gestes métaphoriques) et de mouvements de tête, ce qui semble aller de pair avec le rôle pragmatique des MD qui indique la prise en compte par le locuteur de la dimension interlocutive, comme en témoigne l'exemple 1 dans lequel la deuxième particule fonctionne comme un MD et non pas comme une locution.

Ex 1 : « tu sais tout ce qui était Provence et tout *tout ça* »

**VALEURS DES PARTICULES D'EXTENSION :** Pour chaque particule d'extension, nous avons

<sup>10</sup> Sous Praat. L'annotation de la vidéo, quant à elle, a été réalisée avec Anvil. Les tests de proportion que nous ne mentionnons pas dans ce résumé ont été effectués avec le logiciel R.

<sup>11</sup> Allongement perçu supérieur à l'allongement final de groupe.

annoté une valeur, en reprenant la typologie proposée par Overstreet (2000). Cette annotation a été réalisée sur des critères sémantiques à partir du script uniquement. Les trois valeurs étaient les suivantes :

LIST : liste (ex : « ceux qui font les courses ceux qui font la vaisselle *et cetera* »)

ILL : illustration (ex : « c'est comme les marrons qu'on bouffe *tout ça* c'est des châtaignes aussi »)

INTER : intersubjectivité (ex : « il avait perdu ses parents *tout ça* »)

Pour les catégories fonctionnelles de la prosodie, nous avons extrait des fichiers son sous Praat les particules et leur contexte, puis avons filtré la voix de façon à désémantiser la parole et ne conserver que la mélodie. Nous avons ainsi constitué 6 fichiers WAV de parole filtrée distincts et des textgrids spécifiques dans lesquels étaient uniquement notées les particules (pour conserver l'information sur leur emplacement et leur durée). Puis nous avons noté pour chaque particule 3 fonctions prosodiques, le but étant de vérifier si à chaque valeur sémantique intuitive correspond un contour prosodique spécifique :

CONT Contour de continuation

ENUM Contour d'énumération

T Contour terminal

Une plus grande finesse des contours ne nous a pas semblé réalisable sur de la parole filtrée.

Il apparaît, au vu des résultats, que l'on peut associer le contour ENUM avec la valeur LIST, et que ce contour est réalisé avec une grande variété de mouvements mélodiques et de plages intonatives, mais que sa principale caractéristique est la présence d'un allongement dans le contexte antérieur immédiat, et/ou sur la particule elle-même. Les deux particules les plus susceptibles de marquer une énumération sont « et tout ça » et « et tout », ce qui est congruent avec le fait que ces deux particules sont les moins susceptibles d'être employées comme MD, puisque l'extension de liste est la valeur la plus proche de la signification d'origine et implique une portée syntaxique plus réduite.

En revanche, on observe aussi de nombreux gestes produits sur les particules ayant cette valeur, ce qu'il faudra approfondir.

On peut également associer clairement le contour Terminal avec la valeur Intersubjective. Ce contour est régulièrement situé en plage intonative moyenne, mais les mouvements mélodiques sont également très variés car il n'y a pas plus de contours descendants ou plats pour cette fonction. En revanche, cette avec cette fonction que l'on trouve plus d'occurrences inaccentuées. Là encore, les résultats sont intéressants car si l'on opère une gradation des valeurs et qu'on la met en relation avec le statut MD/Locution, la valeur Intersubjective a une fonction beaucoup plus proche de la fonction méta-discursive des MD que les valeurs LIST et ILL.

Enfin, le contour CONT est le moins spécifié en termes prosodiques et ne peut être associé à une valeur spécifique (il n'est pas associé typiquement à ILL). Ce contour est également réalisé en plage moyenne et de préférence avec des mouvements mélodiques montants. On rencontre également moins de gestes avec ce type de contour.

En résumé, ces premiers résultats sur le fonctionnement et la forme des particules d'extension du français, qui ont une fois encore fait l'objet de peu d'études jusqu'à présent, nous semblent assez encourageants. On voit se dessiner une gradation allant d'une signification d'origine

vers une fonction pragmatique située entre les ponctuants et les *question-tags* anglais. Il faudra cependant sans doute repenser la valeur ILL qui semble pour l'instant n'avoir qu'une réalité sémantique intuitive, non spécifiée sur le plan de la prosodie et de la gestualité. De plus, cette étude a été réalisée sur les quatre particules d'extension les plus fréquentes du français (avec de grandes variabilités dans la fréquence d'emploi de chacune d'elles selon les locuteurs), et il serait intéressant d'étendre l'analyse aux autres particules d'extension. Enfin, cette analyse qui décrit les caractéristiques acoustiques des particules permet d'enrichir la théorie du phrasé prosodique.

### Références

- Bertrand, R. et al. (2008) Le CID - Corpus of Interactional Data - Annotation et Exploitation Multimodale de Parole Conversationnelle. *Traitement Automatique des Langues*, vol. 49, no. 3, p. 1-30.
- Corpus Beeching*, <http://www.uwe.ac.uk/hlss/llas/iclru/corpus.pdf>, enregistrements d'entretiens en français dont la transcription est consultable en ligne.
- Corpus CLAPI*, <http://clapi.univ-lyon2.fr/>, *Corpus de LANGUE Parlée en Interaction* enregistré en situation réelle, dans des contextes variés, développé par le Groupe ICOR, 30h de transcriptions interrogeables par concordancier.
- Corpus Orléans*, <http://bach.arts.kuleuven.be/elicop/>, enregistrements d'entretiens réalisés dans 40 villes de France, repris dans le projet ELICOP (*Etude Linguistique de la Communication Parlée*), 80h de transcriptions interrogeables par concordancier.
- Fraser, B. (1999) "What are discourse markers?" *Journal of Pragmatics* 31: 931-952.
- Lacheret, A., (2004), Structure communicative et géométrie intonative : que nous dit la synthèse de la parole ? *Cahiers de l'Institut de Linguistique de Louvain*, 30 (1-3), pp. 115-139.
- Overstreet, M., (1999). *Whales, candlelight, and stuff like that: General extenders in English discourse*. New York; Oxford: Oxford University Press.
- McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. Chicago and London: The University of Chicago Press.
- Schiffrin, D. (1987). *Discourse Markers*. Cambridge: C.U.P.
- Selkirk E. (2003). Sentence Phonology, *International Encyclopedia of Linguistics*, 2nd ed. Oxford:Oxford University Press.