

Les unités rythmiques et macro-structures à l'oral: délimitation et composition

Résumé

L'organisation de la macro-unité, notamment de l'Episode, est analysé dans le cadre d'une narration à l'oral français spontané. Les phases y sont comparables à celles de Labov et Waletzky (1966) avec un foisonnement de ponctuants, d'évaluations et commentaires de la part du co-énonciateur. Des signaux d'ouverture et de clôture délimitent cette macro-structure. La prise en compte des résultats d'un test de perception ainsi que l'observation d'un corpus composé de discours et d'échanges produits par des locuteurs différents montre que cette unité est rythmique: différentes séquences la composent.

1. Introduction

L'existence de macro-unités et de macro-structures régissant les textes, et notamment le récit est attestée par de nombreuses études, aussi bien en analyse du discours qu'en psycholinguistique. Certaines unités sont envisagées par les psycholinguistes et les fonctionnalistes, il s'agit notamment des concepts de script, de schéma et de frame (Lambrecht)¹ désignant des structures régulières qui servent à décrire respectivement le récit et les textes narratifs, une scène visuelle ou encore une rubrique.

Minsky (1975) désigne par exemple par «frame» un «cadre de connaissances». Le «script»² (Schank et Abelson, 1977) est une unité très utilisée par les psycholinguistes et en Intelligence Artificielle, et est orientée vers la compréhension et la production des récits. Ces structures ou unités ont comme point commun de décrire des situations spécifiques ainsi que les connaissances d'arrière-plan.³ Elles permettent de tenir compte d'un ensemble d'actions, ou d'images etc.... Il s'agit de macro-unités. Mais ces concepts très utilisés à la fois en Intelligence Artificielle mais aussi en analyse de discours (textuelle) expriment surtout soit l'idée de macro-structure (cf. «schéma»), soit l'idée d'arrière-plan (cf. «frame, script»).

Les macro-unités englobent des unités informationnelles (cf. par exemple, la période de Lacheret-Dujour et Victorri, 2002, le paragraphe de Morel et Danon-Boileau, 1998). Par ailleurs, les travaux de Labov et Waletzky (1966) sur les propriétés formelles des narrations ont montré que ces macro-structures sont également composées par des phases. Dans cet article, l'étude d'un corpus et un test de perception ont pour but d'analyser les macro-unités à l'oral français spontané et d'en déterminer le contour et la composition. Nous nous sommes

¹ Caron (1989) l'appelle «cadre de connaissance».

² Gundel *et al.* (2000: 96) dans le «statut cognitif et forme des anaphoriques» définissent le script comme «une séquence prédéterminée d'actions remplies par des acteurs stéréotypiques et comportant une situation bien connue».

basée pour cette étude sur le cadre théorique de *Grammaire de l'intonation*, de Morel et Danon-Boileau (1998); le paragraphe oral est selon eux composé d'un préambule, d'un rhème et d'un post-rhème

2. Corpus et méthodologie

Le corpus est composé d'une série d'enregistrements recueillis entre 2000 et 2004, d'une durée de plusieurs heures, mais dont nous avons choisi d'annoter et de transcrire 70 minutes. Trois locuteurs se partagent ces temps de paroles: Nadia, Benoît et Diane. Le choix de ces trois locuteurs est sociolinguistiquement motivé: en effet, tous les trois travaillent dans le milieu audio-visuel (au sein d'une chaîne de télévision ou de groupes musicaux): ils ont environ le même âge (la trentaine) et appartiennent à la même catégorie socio-culturelle. Notons enfin que les trois locuteurs ont été enregistrés dans différentes situations: entre amis, en famille et dans un environnement professionnel.

Un second corpus a été constitué pour les besoins d'un test de perception: il s'agit de six petits extraits, dont la durée moyenne est de 30 secondes, chaque deux extraits appartenant à l'un de nos trois locuteurs.

Nous avons fait le choix de transcrire méthodiquement 78 minutes de ce corpus correspondant à en moyenne 25 minutes par locuteur choisis à partir d'une écoute auditive. Nous avons analysé trois minutes complètes avec Praat⁴ ainsi qu'un certain nombre de passages, afin d'intégrer l'analyse intonative à notre étude. Nous avons également utilisé le logiciel de statistiques textuelles Lexico3 pour faire des comptages et relever les fréquences significatives de certaines marques.

Une expérience a été conduite sur le second corpus. Il sera détaillé dans la Section 3.1. Une analyse statistique a été entreprise, à partir du logiciel Statview, afin de s'assurer de l'homogénéité des résultats.

Nous avons par ailleurs comparé les résultats du test de perception aux unités informationnelles de Morel et Danon-Boileau (1998) (soit au paragraphe oral constitué de préambule, de rhème et de post-rhème).

3. Analyse

Pour rendre compte de l'existence à l'oral de telles unités, nous avons fait passer un test de perception à des sujets «tout-venants», permettant de saisir la nature des unités «de petite fin» (correspondant à l'écrit à la phrase par exemple) et de «grande fin» (correspondant à l'écrit à

³ Cf. Gundel *et al.* (2000: 99).

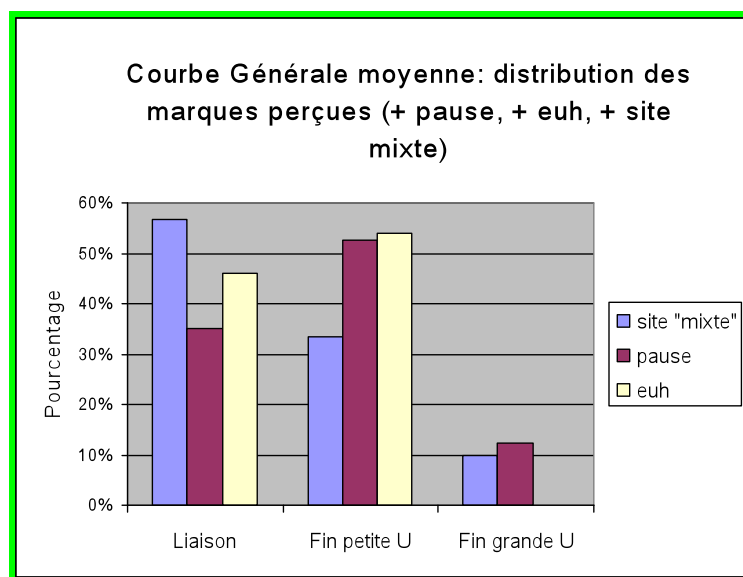
⁴ Il s'agit d'un logiciel d'analyse prosodique.

la fin du paragraphe. Il s'agit dans cet article de d'examiner les unités de «grande fin» et de les délimiter et d'analyser leur composition.

Dans notre test, d'une durée de 1h30 en moyenne, présenté à deux groupes constitués respectivement de 22 et de 23 sujets, nous avons présenté 6 extraits aux sujets, en ordre inversé,⁵ en leur demandant d'abord de «faire une croix là où ils perçoivent une marque de cohésion (marque de liaison, de structuration, d'organisation de l'ensemble)», et ensuite de «tracer un trait oblique chaque fois qu'ils perçoivent une fin de petites unités (1 trait) ou de grandes unités (2 traits)». Les résultats de ce test montrent les caractéristiques perçues syntaxiquement et intonativement des marques de fin de «petites unités» et de fin de «grandes unités».

Le contexte prosodique semble avoir beaucoup joué dans la perception des marques. En effet, nous avons remarqué que les réponses des auditeurs ont été très nombreuses dans le voisinage de pause, de «euh» ou de la pause et du « euh » juxtaposés (ce qu'on a appelé «site mixte» en reprenant ce mot à Candéa, 2000). Ces sites mixtes ont influencé à des degrés divers les réponses.

Figure 1. Courbe générale moyenne: distribution des marques perçues (+ pause, + euh, + site mixte)



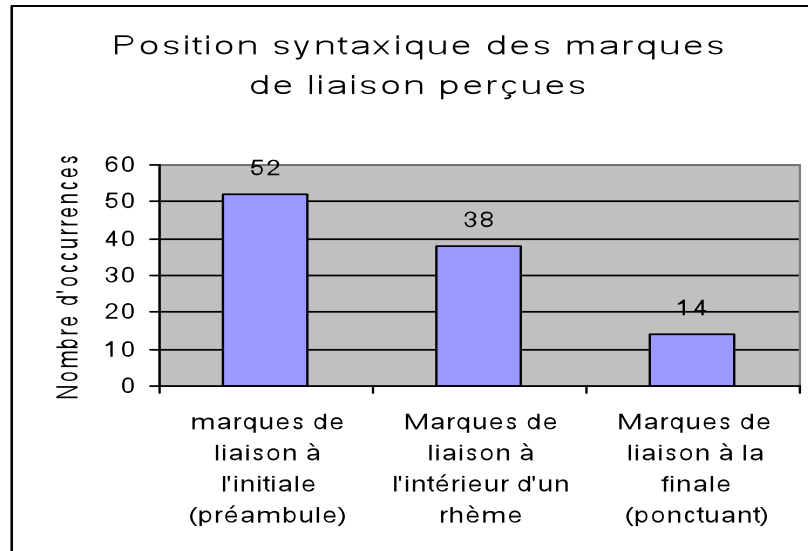
Les marques perçues de liaison se trouvent dans 57% des cas dans le voisinage de sites «mixtes»; les marques perçues de fins de petite unité le sont plutôt à proximité de «euh»: 54% et de pause: 53%. Les marques de fin de grande unité perçues par les auditeurs sont situées

⁵ L'ordre de ces extraits a été inversé pour le groupe 2, dans le but de vérifier si certaines réponses étaient dues au hasard.

avant des pauses dans 12% des cas. Ces chiffres nous donnent une représentation du contexte prosodique dans lequel les marques de liaison et de fin d'unité ont été perçues et notées.

Observons enfin le contexte syntaxique et énonciatif dans lequel se trouvent les marques.

Figure 2. Position syntaxique des marques de liaison perçues



Ce tableau nous montre que 50% des marques de liaison perçues ont été notées à l'initiale (en même temps qu'un site mixte) ce qui correspond au démarrage d'un nouvel énoncé.

Nous avons confronté ces résultats avec le découpage en unités énonciatives selon la théorie de Morel et Danon-Boileau (1998). La position initiale occupée par les marques de continuation fait partie du préambule, première unité dans la théorie du paragraphe oral de Morel et Danon-Boileau (1994).

Plus généralement, les résultats du test de perception montrent que:

Les fins de petite unité ont été perçues aux frontières, c'est-à-dire au début et à la fin de constituant morpho-syntaxique i.e. (préambule, rhème), deux constituants discursifs de la théorie énonciative de Morel et Danon-Boileau (1998).

Les fins de grandes unités ont été perçues dans les voisinages de sites mixtes (interruption de parole et hésitation) ainsi qu'après une pause. Ce sont donc des fins «marquées» qui ont été notées.

22 unités signalées comme marque de fin et suivies par des connecteurs perçus à leur tour comme des unités de liaison ont attiré notre attention. Nous nous sommes alors intéressée aux contextes dans lesquels se trouvent ces unités ainsi qu'à leur position. Ces unités se situent souvent après un rhème, et sont suivies dans 48% des cas par des pauses normales,⁶ et dans

⁶ Nous définissons les mots ci-dessous de la manière suivante :

26% par des pauses moyennes. Nous avons par ailleurs relevé des durées quasi-cycliques entre ces unités. La distance entre les différentes unités qui se succèdent est le plus souvent de 5 secondes dans 66% des cas, de 10 secondes dans 27% des cas et enfin de 15 secondes dans 5% des cas. Nous en avons déduit qu'il s'agit d'une unité rythmique, d'un seuil de perception particulier correspondant à une unité qui peut être plus large que le paragraphe oral. Ce qui nous a conduit à envisager l'existence de nouvelles unités orales, autres que celle du paragraphe intonatif. En effet, l'existence d'unités rythmique à l'oral est avéré par exemple par des études portant sur les échanges mère-enfant: les bébés seraient sensibles aux frontières de phrases (linguistique ou musicales s'appuyant sur «l'information acoustique inhérente aux fins de phrases, comme la modification du contour mélodique ou l'allongement de la durée du son (Fassbender, 1995)».⁷ Le bébé a donc une capacité à détecter une «macro-structure rythmique» grâce à son contour mélodique.

Les résultats du test nous ont poussé à observer de près la structure des unités «de petite et grande fin» à l'oral. L'examen du corpus de 78 minutes a montré l'existence de macro-structures (particulièrement dans les narrations), de matrices relativement stables.

Cette matrice rejoint les résultats du test de perception notamment en ce qui concerne les durées cycliques de certaines parties de la macro-unité: elles sont régulières et peuvent durer 7 secondes, 15 secondes et jusqu'à 30 secondes: il y a donc bel et bien un rythme à prendre en considération au sein de ces unités.

Par ailleurs, les démarcations initiales et finales sont opérées par des ponctuants⁸ et ligateurs,⁹ conformément aux résultats du test.

4. Résultats: l'Episode

Le terme d'épisode a été utilisé par de nombreux auteurs, par exemple Van Dijk (1978), pour rendre compte de la structure du récit. C'est un des constituants hiérarchiquement organisés de la grammaire des récits (cf. Mandler et Johnson (1977), voir Mandler (1982) pour un bilan sur ces grammaires).

Selon Caron (1989: 220), les épisodes seraient «des unités de traitement»:

-
- la pause normale: les pauses ont une durée normale lorsqu'elles se situent entre 20 et 40 centisecondes (cf. Morel et Danon-Boileau, 1998; Candea, 2000):
 - la pause moyenne est entre 60-80 centisecondes,
 - la pause longue se situe au delà de 100 centisecondes (cf. Candea, 2000).

⁷ Voir Devouche et Gratier (2001: 60).

⁸ Il s'agit «d'un marqueur spécialisé» qui termine un segment: le rhème; il peut également clore le paragraphe (Morel et Danon-Boileau, 1998).

⁹ Il s'agit du premier constituant du préambule et «précise le lien de ce qui va se dire avec ce qui l'a déjà été» (Morel et Danon-Boileau (1998: 39).

«ils constituent des «chunks» en mémoire (Black et Bower, 1979); en outre, le temps de traitement est significativement plus élevé au début et à la fin d'un épisode (Haberlandt, Berian et Sandson, 1980): cette différence traduirait l'initialisation d'une nouvelle unité en mémoire au début de l'épisode, et une organisation en «macro-unité» à la fin; ce qui rapprocherait le traitement des unités narratives de celui du traitement des propositions syntaxiques».

Caron désigne par épisode un «constituant du récit»; nous désignons par épisode une macro-unité de l'oral, composée de plusieurs paragraphes oraux. Il est par ailleurs constitué de phases, dont la démarcation est assurée par la présence de ligateurs, de supports lexicaux disjoints¹⁰ et cadres,¹¹ mais aussi de ponctuations propres à l'oral, parmi lesquels nous incluons les marques d'écoute de l'interlocuteur. Ces derniers permettent également de délimiter les différentes séquences (cf. ci-après) de l'épisode. Nous définissons les séquences comme des unités d'action ou d'événements qui peuvent être narratives, descriptives ... La séquence se caractérise par une unité énonciative (pronoms et marqueurs de coréférence), par une unité temporelle (même temps) et textuelle (unité argumentative). Nous verrons des exemples de séquences dans l'épisode ci-dessous.

4.1 Définition et propriétés

Nous nous référons aux définitions de l'épisode fournies par les psycholinguistes Kintsch et Van Dijk (1978); Mandler (1982); Black et Bower (1979); Haberlandt, Bérian et Sansdon (1980); (cf. Caron, 2001) et notamment au concept de schéma. En effet, le schéma comporte une structure narrative (Caron, 2001; Fayol, 1985: chapitre IV): selon les grammaires de récits, le récit est composé d'une exposition ainsi que d'une structure événementielle. La structure événementielle est composée d'épisodes (x N) qui se succèdent. L'épisode est lui-même structuré comme suit:

Episode	→Début-Développement-Fin
Développement	→Réaction simple-action →Réaction complexe-chemin vers un but.

¹⁰ La définition donnée par Morel et Danon-Boileau précise que:

«Un dernier segment, le support lexical disjoint, vient parfois mettre en place la référence de l'argument qui sert de support à la prédication du rhème. La plupart du temps, dans l'oral spontané du français, ce support se trouve dissocié intonativement et syntaxiquement du rhème. Il est alors rappelé auprès de ce dernier par un indice pronominal (pronom personnel, démonstratif ou relatif, cf. Chapitres 4 et 9).» (Morel et Danon-Boileau, 1998: 41).

¹¹ Le cadre permet dans l'approche co-énonciative de Morel et Danon-Boileau de délimiter des zones thématiques relatives aux circonstances de l'énonciation.

Nous donnons ci-dessous un exemple d'épisode, composé de 2 séquences, de phases de durées variables :

Episode

Séquence 1 (20 secondes)

❶ Entrée en Matière (3 secondes)

en fait
c'est c'est assez drôle ↗
pa(r)c' que

❷ Description (10 secondes)

t'as tous les vieux ↗ euh Durand ↗ °Durand il était
juste à côté de nous ↗ euh tu vois
Guillaume Durand ° ↗
-§mm§
et euh

les vieux euh les vieux has been comme ça
qui ont été qui sont mis au placard euh
temporairement ↗ euh

❸ Commentaires (7 secondes)

ils sont {}
donc eux
ils sont hyper aigris ↗
ils sont là ↗
ils se font discrets ↗
ils bougent pas et tout ↗

-§ouais§ -----

Séquence 2 (15 secondes)

❹ et ❺ Rupture et description contrastée (5 secondes)

-{} puis en fait ↗ euh
toutes les équipes des chaînes
bah

qui sont comme nous ↗
qui sont souvent ↗
°bon un peu plus vieux que nous quand même°
mais euh ↘
-§ouais §

❻ Commentaires (8 secondes)

-mais
jeunes ↗ dynamiques ↗
-§ouais§
-qui ont des idées ↗
qui sont créatifs ↗
qui bricolent euh quasiment tous autant que nous ↗ °tu vois°
qui font en sorte de faire une belle chaîne de télé ↗
où là eux ils mettent un peu le feu °quoi° ↘

❼ et ❽ Manifestation de l'écouteur et chute (2 secondes)

-§ouais §

Cet épisode comprend 6 phases, réparties de la façon suivante: 1-3-4 (répétées chacune)-5, suivies des phases 5-6.

Par ailleurs, l'épisode comprend un certain nombre de traits:

- Des signaux d'ouverture: l'initialisation de l'épisode est marquée par une formule générique «c'est assez drôle» précédée de «en fait» (un commentaire modal et un marqueur de contraste, de point de vue). Notons que l'épisode et les séquences narratives

peuvent être initialisés par des présentatifs tels que «il y a» ou «c'est» qui jouent aussi un rôle cadratif à l'oral.

- Des signaux de clôture ou de fermeture: la chute est souvent marquée soit par une manifestation sonore de l'écouteur §mm§, §ouais§ ou par un ponctuant en intonation montante ou descendante «et tout».

Lors de la passation du test de perception, 38% des sujets ont perçu comme marque de fin un ponctuant ou marqueurs de fin, un syntagme nominal ensuite (33%), une coréférence (10%) ou des adverbes (9,61%). Ce classement fait surtout ressortir à l'oral deux classes: celles des ponctuants mais aussi des post-rhèmes (coréférence, syntagme nominal...). Les ponctuants et les post-rhèmes peuvent jouer un rôle important comme indices de fermeture d'une unité à l'oral. Ce sont des marqueurs de clôture de l'épisode et des séquences.

- La durée des séquences est variable dans l'exemple ci-dessus: de 7 secondes (assez rarement) à 15 secondes et surtout 30 secondes.

Or, dans le test de perception, et sur des petits extraits, des unités cycliques de 5 à 15 secondes ont été relevées par les sujets, considérées comme des sites où se trouvaient «des unités de liaison et de fin». Un rythme se dégage de cette régularité, qui mériterait d'être étudié plus avant.

Ces durées se rapprochent par ailleurs de celles relevées par les chercheurs sur les échanges mère-bébé. En effet, «Trevarthen (1999) a décelé des intervalles temporels longs, de l'ordre d'une trentaine de secondes, au sein de l'interaction vocale mère-bébé, et les a rapprochés d'unités narratives que l'on retrouve dans des chants et poésies de par le monde» (Devouche et Gratier, 2001: 76-77).

5. Conclusion

Dans cet article, nous avons montré l'existence d'unités rythmiques à l'oral au sein de macro-structures: nous proposons de les appeler respectivement séquence et épisode. L'épisode se compose d'un certain nombre de phases, relativement stables et de séquences, unités thématiques et rythmiques qui peuvent durer jusqu'à 30 secondes. L'épisode s'ouvre grâce à des signaux d'ouverture (à travers généralement des présentatifs), et se termine par des signaux de clôture (des ponctuants et des post-rhèmes).

La cohésion interne de l'épisode s'opère grâce à un ensemble de marqueurs, qui appartiennent très souvent à l'interlocuteur ou au co-énonciateur: les réactions de celui-ci tout au long de l'épisode, à travers des ponctuants et des évaluations diverses, constituent des phases. Les séquences et les phases s'imbriquent au sein de la macro-unité selon des modalités diverses: la composition de l'épisode dépend du type de corpus pris en compte. Nous projetons

d'élargir ces analyses à des types d'échanges divers. Par ailleurs, l'étude du contour prosodique de l'Episode fera l'objet d'un autre article.

Références

- Black, J. B. et Bower, G. H. (1979) Episodes as chunks in narrative memory. *Journal of Verbal Learning and Verbal Behavior*, 18: 309-318.
- Candea, M. (2000) Contribution à l'étude des pauses silencieuses et des phénomènes dits «d'hésitation» en français oral spontané. Thèse de doctorat, Université de Paris 3 – Sorbonne Nouvelle.
- Caron, J. (2001) *Précis de psycholinguistique*. Paris: Presses universitaires de France (Quadrige).
- Devouche, E et Gratier, M. (2001) Micro-analyse du rythme dans les échanges vocaux et gestuels entre la mère et son bébé de 10 semaines. *Devenir*, 13: 55-82.
- Fassbender, C. (1995) La sensibilité auditive du nourrisson. In: I. Deliège et J. A. Sloboda (éds.) *Naissance et développement du sens musical*. Paris: Presses universitaires de France.
- Fayol, M. (1985) *Le récit et sa construction*. Genève: Delachaux et Niestlé.
- Gundel, J. K. Hedberg, N. and Zacharski, R. (1993) Cognitive Status and the form of referring expressions in discourse. *Language*, 69: 274-307.
- Haberlandt, K., Berian, C. et Sandson, J. (1980) The episode schema in story processing. *Journal of Verbal Learning and Verbal Behavior*, 19: 635-650.
- Kitsch, W. et Van Dijk, T.A. (1984) Vers un modèle de la compréhension et de la production de textes, *Psychological Review* 1978, 85: 5 (traduction française in G. Denhière 1984).
- Labov, W. et Waletzky, J. (1966) Narrative analysis: oral versions of personal experience, In: J. Helm (éd.), *Essays on the verbal and visual arts: Proceedings of the 1966 Annual Spring Meeting of the American Ethnological Society*, Seattle (University of Washington Press), 12-44.
- Lacheret-Dujour, A. et Victorri, B. (2002) La période intonative comme unité d'analyse pour l'étude du français parlé, modélisation prosodique et enjeux linguistiques. *Verbum*, 24: 55-69.
- Mandler, J. M. (1982) Recherches récentes sur les grammaires de récits. In : J.-F. Le Ny et W. Kitsch (éds.) *Language and comprehension*. Amsterdam/New York/Oxford: North Holland, 705-715.
- Morel, M.-A. (1994) Valeur énonciative des variations de hauteur mélodique en français, *Journal of French Language Studies*, 5: 189-202.
- Morel, M.-A. et Danon-Boileau, L. (1998) *Grammaire de l'intonation*. Paris: Ophrys.
- Morel, M.-A. (2004) Intonation, Regard et genres dans le dialogue à bâtons rompus, *Langages*, 153: 15-27.
- Minsky, H. P. (1975) *John Maynard Keynes*. New York: Columbia University Press.
- Schank, R. and Abelson, R. (1977), *Scripts, plans, goals and understanding*. Hillsdale, N.J.: Lawrence Erlbaum.
- Trevarthen, C. (1999) Musicality and the intrinsic motive pulse. *Musicae Scientiae, Special Issue*, 155-215.
- Van Dijk, T.A. (1978) *Tekstwetenschap*. Utrecht: Spectrum.

Conventions de transcription

§§ note la fin d'un chevauchement de paroles

e/ euh note ce qui est généralement transcrit ‘euh’ dans la graphie traditionnelle

: note l’allongement d’un son ; plus il y a de points, plus l’allongement est long
(: / :: / ::: /)

MAJ indiquent une prononciation appuyée : accent d’intensité, d’insistance

_{Indice} indique une incise ou une intonation basse et plate

° ° notent des fragments prononcés sur un ton plus bas que les paroles qui les encadrent
(incises basses)

m son prononcé lèvres closes correspondant le plus souvent à une marque d’approbation
et/ou de confirmation du discours de l’interlocuteur

{30} durée des pauses silencieuses en centisecondes

H souffle

↗ fréquence fondamentale (F0) montante

↘ fréquence fondamentale (F0) descendante.