



Revue MÉTHODAL

REVUE INTERNATIONALE DE MÉTHODOLOGIE DE L'ENSEIGNEMENT/APPRENTISSAGE DES LANGUES

**Cet article est extrait de
Méthodal, n° 6 (2026)**

**Actes du colloque international
*Quelles « grammaires enseignantes » pour l'enseignement des langues ?***

9–10 octobre 2024

Sous la direction de
Céline Peigné · Jean-Claude Beacco · Béatrice Blin

Ce document constitue un tiré à part numérique d'un article publié dans ce numéro de Méthodal.

Numéro complet : ISSN 2585-2892 · <https://revue.methodal.net> ↗

© Les auteurs · Licence Creative Commons BY-NC-SA

Des manuels aux corpus : Comment décrire le suffixe *-Dir* en turc en contexte didactique ?

Sülün AYKURT-BUCHWALTER

Université Paris Nanterre, MoDyCo UMR 7114, France

Mots-clés

ENSEIGNEMENT DE LA GRAMMAIRE – TURC L2 – MÉTHODES DE LANGUE – ANALYSE DE
CORPUS – MARQUEURS MODAUX

Keywords

GRAMMAR TEACHING – TURKISH AS A SECOND LANGUAGE – LANGUAGE TEXTBOOKS – CORPUS
ANALYSIS – MODAL MARKERS

Anahtar kelimeler

GRAMER ÖĞRETİMİ – TÜRKÇE İKİNCİ DİL – DİL ÖĞRETİMİ DERS KİTAPLARI – DERLEM ANALİZİ – KIPLIK
BELİRTEÇLERİ

Résumé

Le suffixe *-Dir* constitue un élément de la langue turque difficile à décrire et, par conséquent, difficile à enseigner et à acquérir en L2. Parfois présenté comme un marqueur de 3^e personne, il est plus généralement considéré comme un marqueur modal ambigu pouvant exprimer la certitude, ou à l'inverse, l'incertitude. Cette étude a pour objectif de comparer ses différentes descriptions dans du matériel pédagogique de turc L2, et de mettre en perspective ces descriptions par rapport à un grand corpus du turc. Des pistes pour l'exploitation de ce type de données par les enseignants de turc L2 sont proposées.

Abstract

One element of the Turkish language is difficult to describe, and therefore difficult to teach and acquire in L2: the *-Dir* suffix. Sometimes presented as a 3rd person marker, it is more generally considered to be an ambiguous modal marker that can express certainty, or conversely, uncertainty. This study aims at comparing its various descriptions in L2 Turkish teaching material, and putting these descriptions into perspective in relation to a large corpus of Turkish. Suggestions are made for the use of this type of data by teachers of L2 Turkish.

Özet

-Dir son ekinin Türkçede tanımlanması oldukça güçtür, ve aynı zamanda, yabancı dil Türkçe öğretim ve öğreniminde zorluklara yol açmaktadır. Bazı kaynaklarda 3. tekil şahıs belirteci olarak sunulsa da, genellikle kesinlik ya da aksine belirsizlik ifade edebilen, anlamı değişkenlik gösterebilen bir kiplik

unsuru olarak öne çıkabilmektedir. Bu çalışmanın amacı, yabancı dil Türkçe öğretim materyallerinde -Dİr'a ilişkin olarak yer alan farklı tanımlamaları karşılaştırmak ve bu tanımlamaları bir Türkçe derlemiyle karşılaştırarak değerlendirmektir. Yabancı dil Türkçe öğretmenlerinin bu tür verileri kullanabilmeleri için öneriler sunulmaktadır.

Introduction

Les descriptions du suffixe *-Dİr* dans les ouvrages destinés aux apprenants du turc langue étrangère (L2) sont très variées. Parfois décrit comme un simple marqueur de troisième personne, il peut être introduit dès les niveaux débutants (Bozdémir, 2023). Cependant, il est souvent désigné comme un marqueur modal ambigu devant être interprété en fonction du contexte (Ketrez, 2012). Certains ouvrages mettent davantage l'accent sur la valeur de supposition de ce suffixe (Lewis, 2000), tandis que d'autres insistent sur sa valeur d'assertivité (Underhill, 2003). Une telle ambiguïté est susceptible de rendre l'enseignement/apprentissage de ce suffixe particulièrement complexe. De plus, il est légitime de s'interroger sur les sources d'information des auteurs de ces ouvrages lorsqu'ils évoquent la fréquence ou la rareté de telle ou telle valeur du suffixe *-Dİr* dans la langue. Une approche fondée sur les corpus pourrait permettre d'apporter quelques réponses à ces questionnements. Des données issues du langage authentique et en contexte sont susceptibles d'informer les enseignants et les apprenants sur les valeurs *-Dİr*, sans dépendre exclusivement d'exemples « fabriqués », ainsi que sur la fréquence de ces différentes valeurs, au-delà des descriptions proposées par les manuels.

Ainsi, comment une mise en perspective des descriptions fournies dans des manuels de grammaire et des données issues d'un grand corpus de la langue turque pourrait-elle apporter un éclairage sur la description la plus adaptée de *-Dİr* en contexte didactique ? Quelles utilisations par les enseignants pourraient être envisagées ?

1. Cadrage théorique

1.1. Le suffixe *-Dir* en turc moderne : copule ou marqueur modal ?

Le turc est une langue agglutinante et les suffixes y jouent un rôle essentiel : ils sont « en nombre assez limité, mais [leurs] combinaisons multiples permettent d'exprimer des notions très variées » (Bazin, 1987[2000] : 17). La forme d'un même suffixe varie en raison des règles de phonétique (harmonie vocalique et consonantique, Goldstein, 1999 : 31) ; c'est pourquoi *-Dir*, qui appartient à la classe des suffixes de conjugaison, peut prendre les formes *-dur*, *-dür*, *-tir*, *-tir*, *-tur* et *-tür*.

La première information dont il faut tenir compte pour cerner le suffixe *-Dir* est l'absence, en turc, de verbe exprimant « être » au temps présent : ce sont des suffixes de personne qui accomplissent cette fonction. Comme nous le voyons dans l'exemple (1), les première et deuxième personnes portent une terminaison, mais il est tout à fait possible d'exprimer le verbe être à la troisième personne du singulier sans aucun suffixe :

(1)

<i>doktorum</i>	= je suis médecin
<i>doktorsun</i>	= tu es médecin
<i>doktor</i>	= il est médecin

Pourtant, l'apprenant du turc découvre rapidement l'existence du suffixe *-Dir*, souvent présenté comme un suffixe facultatif de la troisième personne du singulier pour exprimer le verbe être. Pour tenter de clarifier ce problème et commencer à expliciter le sens du suffixe *-Dir*, nous proposons un premier regard sur sa description dans les grammaires.

Dans les grammaires du turc langue maternelle, il n'y a pas de consensus sur la dénomination du suffixe *-Dir*. Le terme de *bildirme eki* (« suffixe d'information / de déclaration ») est employé dans certains ouvrages, tandis que *ek-fiil* (« verbe-suffixe »), *cevher fiili* (« verbe substantif ») ou *koşaç* (« copule ») apparaissent dans d'autres ; ces variations dans la terminologie sont susceptibles de conduire à des difficultés pour les enseignants et les apprenants du turc langue maternelle et étrangère (Zülfikar, 2017).

Des travaux en diachronie montrent qu'il s'agit de la forme dérivée du verbe auxiliaire du turc ancien « *tur-* » combiné au suffixe aoriste « *-ur* », exprimant à l'origine le sens de « se tenir » ou d' « exister ». À travers les siècles, la fonction de *-Dir* aurait évolué, renforçant le degré de certitude du propos, puis pouvant également exprimer la probabilité (Ekmekçi Aşan, 2021).

En ce qui concerne le turc moderne, Korkmaz (2003), employant la dénomination de « *bildirme eki* » (« suffixe d'information / de déclaration »), attribue deux fonctions à *-Dir* : l'une correspond en effet à la terminaison de la troisième personne pour exprimer le verbe être ; l'autre exprime une valeur modale. Elle classe les deux fonctions de la manière suivante :

(i) Lorsqu'il est suffixé à un nom ou un adjectif, à la troisième personne, *-Dir* sert à transformer ces derniers en verbe, exprimant ainsi le sens du verbe « être » (p.703), ce qui correspond à la notion de copule :

(2)

1 ^e p. s.	<i>Yalnız-ım</i>	je suis seul(e)
2 ^e p. s.	<i>Yalnız-sın</i>	tu es seul(e)
3 ^e p. s.	<i>Yalnız-dir</i>	il/elle est seul(e)

(ii) Lorsqu'il est suffixé à un verbe, il peut (p. 727) :

- Exprimer la certitude : lorsqu'il est ajouté aux conjugaisons avec le présent duratif « *-mAktA* », le passé non-constaté « *-mİş* » ou le futur, à la troisième personne. Pour cette fonction, de nombreux exemples issus de récits historiques et de romans sont proposés, tels que :

(3)

<i>Halk</i>	<i>Selçukluyu</i>	<i>suçlamaktadır.</i>
le peuple	les seldjoukides + accusatif	accuser + présent duratif + dir ¹

Le peuple accuse les Seldjoukides.

- Exprimer la supposition : à toutes les personnes, lorsqu'il est suffixé à certains modes (p.728).

(4)

<i>Bu</i>	<i>yöreye</i>	<i>herhalde</i>	<i>daha</i>	<i>önce</i>	<i>de</i>	<i>gelmişindir</i>
cette	région + directif	sûrement	plus	avant	aussi	venir + passé non constaté + 2 ^e p.s. + dir

Tu es certainement déjà venu dans cette région.

Ainsi, dans cette description, la fonction de copule existe, mais *-Dir* est également porteur de deux valeurs modales potentiellement opposées. L'indice proposé pour identifier la valeur modale de *-Dir* est le contexte morphologique immédiat, c'est-à-dire le temps du verbe auquel *-Dir* est suffixé ; pourtant, dans l'exemple (4), nous observons la présence d'un adverbe modal (*herhalde*, « probablement, sûrement »), qui constitue un indice pour l'interprétation. De plus, la fonction modale est exclue pour la catégorie (i), à savoir lorsque *-Dir* est suffixé à un nom.

Tura Sansa (1986) rejette cependant l'idée que *-Dir* puisse être un simple marqueur de troisième personne et désigne ce suffixe comme un « opérateur sur l'échelle modale » (p.145) allant du certain au non-certain, dont le sens doit être interprété grâce à des indices contextuels, les temps marqués dans le verbe ou encore la présence de certains marqueurs de modalité. L'auteure souligne que *-Dir* ne suit jamais l'aoriste, car ces deux éléments ont des fonctions sémantiques « totalement parallèles » en ce qui concerne les formes verbales.

¹ A travers les exemples fournis à partir des ouvrages cités, le glosage est de nous ; la traduction est de l'auteur si elle est fournie, et de nous si ce n'est pas le cas. Les abréviations utilisées dans le glosage sont p.s. pour « personne du singulier » et p.p. pour « personne du pluriel ».

La diversité des valeurs que *-Dir* peut exprimer, voire les contradictions dans les valeurs modales sont telles que son enseignement / apprentissage peuvent poser des difficultés aux enseignants comme aux apprenants du turc L2. Des études montrent en effet que l'acquisition de certains affixes en langue étrangère (L2) est rendue plus difficile par leur caractère polyfonctionnel (Mochizuki & Aizawa, 2000).

1.2. L'exploitation pédagogique des corpus en L2

Lorsque la description d'une forme linguistique en langue cible est ambiguë et que sa compréhension nécessite une interprétation du contexte, il est utile pour l'apprenant de découvrir cette forme dans des usages authentiques de la langue. Les corpus, généralement définis comme un ensemble numérisé de textes authentiques écrits ou oraux représentatifs d'une langue ou d'une variété de langue (McEnery & al., 2006), peuvent être exploités dans les processus d'apprentissage : il s'agit du *data-driven learning* (Johns, 1991) ou apprentissage sur corpus (Boulton, 2024). En ayant accès aux corpus dans une approche directe, dans laquelle l'apprenant manipule directement les corpus, ou indirecte, autrement dit, par l'intermédiaire d'une sélection opérée en amont par l'enseignant (Yoon & Jo, 2014), l'apprenant peut découvrir par lui-même et de manière implicite le fonctionnement de la langue cible, dans un esprit d'autonomie et de découverte (Cheng, 2010). En cela, certains chercheurs considèrent que l'exploitation des corpus rapproche quelque peu le processus d'apprentissage de la L2 de l'acquisition de la L1 (Gilquin, 2021). En didactique des langues étrangères, les corpus ont pu servir à l'élaboration de référentiels et de matériel pédagogique, tels que des dictionnaires d'apprenants. Cependant, les enseignants de L2 sont généralement peu formés à ces approches, ce qui rend difficile leur application en classe de langue. Des dispositifs visant à renforcer la « littératie de corpus » des enseignants de L2 (Zareva, 2017) peuvent être confrontés à des défis, tels que le manque de temps, les contraintes institutionnelles ou encore les réticences des enseignants (Aykurt-Buchwalter, 2023). De plus en plus de chercheurs affirment que les enseignants de L2 devraient être formés, en formation initiale et/ou continue, à des approches sur corpus (Farr & Leńko-Szymańska, 2024 ; Götz & Granger, 2024).

2. Méthodologie

2.1. Analyse de manuels

La démarche méthodologique adoptée pour cette étude s'articule autour de deux axes : d'une part, nous analysons neuf ouvrages qui ont pour point commun de s'adresser à des apprenants du turc L2. Il s'agit de méthodes de langue ou de grammaires pratiques. Nous observons à quel niveau de la progression et de quelle manière le suffixe *-Dir* est présenté. Le tableau 1 reprend les ouvrages analysés.

Titre	Auteurs	Date	Editeur
<i>Introduction à l'étude pratique de la langue turque</i>	Bazin, L.	1987 [2000]	Maisonneuve
<i>Grammaire du Turc. Ouvrage pratique à l'usage des francophones</i>	Golstein	1999	L'Harmattan
<i>Turkish Grammar</i>	Lewis, G.L.	1967 [2001]	Oxford University Press
<i>Turkish Grammar</i>	Underhill, R.	2003	MIT Press
<i>Colloquial Turkish</i>	Aarsen, J. & Backus, A.		Routledge
<i>Méthode de turc Volume I</i>	Bozdémir, M.	2023	L'Asiathèque
<i>Méthode de turc Volume II</i>	Berk, S. & Bozdémir, M.	2014	L'Asiathèque
<i>A Student Grammar of Turkish</i>	Ketrez, N.	2012	Cambridge University Press
<i>Yedi İklim B1</i>		2021	Yunus Emre Enstitüsü

Tableau 1. Ouvrages analysés

Nous proposons de mettre en perspective les descriptions qui se trouvent dans ces ouvrages avec des analyses de corpus permettant de mesurer la fréquence d'une forme donnée, en l'occurrence de *-Dir*, dans la langue, à travers différents contextes d'utilisation.

2.2. Analyses de corpus

L'analyse de corpus est basée sur le *Turkish National Corpus* (TNC, Aksan & al., 2012). Le TNC est accessible aux chercheurs sur demande. Il est composé de 50 millions de mots, dont 98% de textes écrits appartenant à différents genres et couvrant une période de 20 ans (1990-2009) et 2% de transcriptions de discours oraux.

Notre analyse porte une forme unique : « *var+dir* » (il y a + *dir*). Le choix de cette forme se justifie par les motifs suivants :

- « *var+dir* » exprime deux sens différents, ce qui permet d'observer des situations variées : en plus d'être l'équivalent d' « il y a », « *var* » permet de construire l'équivalent du verbe avoir, en combinaison avec le suffixe possessif.
- Le choix de « *vardir* » permet de ne pas choisir entre différents temps et modes pour un verbe conjugué. En neutralisant la variable du contexte morphologique immédiat, nous pouvons nous concentrer sur l'impact d'autres critères susceptibles d'interagir avec le sens de *-Dir*.

L'analyse permet de distinguer différents genres de textes écrits, formels et informels, et d'autre part des genres oraux. Parmi les genres écrits, nous n'avons pas retenu la presse écrite, en raison du grand nombre de citations rendant difficile de faire la part entre discours oral et discours écrit. Les deux sous-corpus oraux retenus sont ceux où la fréquence de *vardir* est la plus élevée. Nous reprenons dans le Tableau 2, parmi les genres répertoriés par le TNC, les sous-corpus que nous avons retenus pour nos analyses.

	Nombre de documents	Nombre de mots
Écrit		
– Écrit scientifique (sciences sociales et humaines, arts)	786	5 497 606
– Écrit non scientifique (sciences sociales et humaines, arts)	243	5 055 382
Écrit informel		
– forum	68	468 038
– blog	119	1 199 927
Oral formel		
– Assemblée nationale	11	83 381
Oral informel		
– conversation	205	426 443

Tableau 2. Sous-corpus du *TNC* retenus pour cette étude

Ainsi, notre étude porte sur un corpus écrit d'environ 12 millions de mots et un corpus oral d'environ 500 000 mots. À travers ces différentes sections du corpus, nous recherchons la fréquence relative des cooccurrences de *vardır* avec des marqueurs modaux exprimant la certitude et l'incertitude, à savoir cinq adverbes situés sur une échelle de certitude (*belki*, *herhalde*, *kesinlike*, *elbette*, *mutlaka* / peut-être, probablement, certainement, bien sûr, absolument), dans un contexte de -5 à +1 mots.² En raison de la structure Sujet-Objet-Verbe du turc, la présence du marqueur modal à gauche du verbe nous semble plus pertinente. Nous complétons ces analyses avec des tests de vraisemblance (*log likelihood*), permettant de mesurer statistiquement le degré d'association entre les occurrences de « *vardır* » et celles des marqueurs modaux.³

² Quelques variations sont possibles concernant les marqueurs modaux recherchés : dans le corpus oral et d'écrit informel, la forme "herhalde" apparaît parfois comme "heralde" ; "kesin" est une variante plus informelle de "kesinlike" (certainement) et "elbet" est un équivalent de "elbette".

³ Les rapports de vraisemblance sont calculés avec l'outil <https://ucrel.lancs.ac.uk/llwizard.html> de l'Université de Lancaster, ou sont fournis par l'interface du TNC lors des recherches de cooccurrences.

3. Résultats

3.1. Le suffixe *-Dir* dans les manuels

L'analyse des manuels montre des similitudes et des divergences concernant la description de *-Dir*. Différents critères apparaissent à travers les manuels pour tenter de décrire les valeurs possibles de ce suffixe, tels que la distinction écrit/oral, le registre de langue et le genre discursif.

Dans la majorité des ouvrages étudiés, *-Dir* est introduit tôt, avec pour objectif d'apprendre à exprimer l'équivalent du verbe « être ». Sachant que « le turc n'a pas de présent du verbe être », Bazin (1987[2000] : 59) indique que « Ali est turc » peut se dire : *Ali, Türk* ou *Ali, Türk'tür*. Cependant, plus loin dans l'ouvrage, l'auteur décline différents cas de figure concernant l'emploi de *-Dir*, en fonction de la distinction oral/écrit, et émet des hypothèses liées à la fréquence des différents usages (p.87) :

- « toujours facultatif et **peu employé** dans la langue parlée »
- après *-lyor*, « **très rare**, et réservé à une forme d'insistance »
- « dans la langue écrite, **habituel** après *-mekte (...)*, *-(y)ecek* et *-meli* »

L'auteur exclut aussi l'usage de *-Dir* avec l'aoriste, sans toutefois souligner les parallélismes sémantiques entre les deux notions.

Dans *Turkish Grammar* (Lewis, 2001), la forme ancienne « *turur* » est mentionnée, et *-Dir* est d'abord présenté comme une copule facultative, autrement dit, comme un suffixe du verbe être à la 3^e personne, laissé entre parenthèses car « *in Turkish, simple A is B equivalences are expressed without a copula* » (p. 96). Toutefois, l'auteur fait référence à la variable du registre (« discours formel / discours ordinaire / informel ») et à la distinction écrit/oral, pour décrire des valeurs bien distinctes (emphase ou supposition) et fait une hypothèse sur la fréquence des valeurs (« *more often* ») :

In writing and in formal speeches -dir expresses the copula. (...) In ordinary speech -dir is not used in such simple A=B sentences; (...) the use of -dir in informal speech is either for emphasis or, more often, to indicate a supposition. (Lewis, 2001 : 97).

Cependant, l'exemple qui illustre l'emploi de *-Dir* est en décalage avec cette description : « *Kızın adı Fatma'dır* » (Le prénom de la fille est Fatma) (p.97), est une phrase A=B et ne relève pas du discours formel. La compréhension par le lecteur / apprenant n'est donc pas facilitée par l'exemple fourni.

On trouve dans cet ouvrage une référence à la diachronie qui semble intéressante : selon Lewis, une des manifestations du rapprochement entre écrit et oral en turc est la disparition progressive de *-Dir* employé comme simple copule à l'écrit. Autrement dit, dans des textes plus anciens, *-Dir* pourrait être une copule, mais dans des textes plus récents, il est un indice de formalité. Cette évolution diachronique n'apparaît pas dans les autres ouvrages.

Goldstein (1999) n'introduit pas *-Dir* dans les conjugaisons équivalentes au verbe « être ». Il affirme plutôt que *-Dir* a deux fonctions possibles : d'une part, il est un marqueur de « langage soutenu », auquel cas il « s'ajoute sans modification de sens à la 3^e personne du singulier » (p.159) de certaines conjugaisons, dont le passé non-constaté, le présent duratif *-mAktA*, et le futur.

(5)

<i>Yakında</i>	<i>parlak</i>	<i>bir</i>	<i>çağ</i>	<i>başlayacaktır.</i>
bientôt	brillant	un	âge	commencer + futur 3 ^e p.s. + tir

Bientôt un âge doré commencera.

D'autre part, *-Dir* peut exprimer une « vaste palette de degrés d'incertitude, depuis peut-être jusqu'à probablement » (p.159).

L'exemple fourni est le suivant :

(6)

<i>Siz</i>	<i>de</i>	<i>bu</i>	<i>akşam</i>	<i>geliyorsunuzdur.</i>
vous	aussi	ce	soir	venir + présent + 2 ^e p.p.+dur

Vous aussi venez **probablement** ce soir.

Dans cet ouvrage, il n'y a pas de référence à la fréquence de telle ou telle fonction ou valeur.

Underhill (2003), tout en introduisant *-Dir* dès le début de son ouvrage comme un « auxiliaire » avec l'exemple de la phrase « *Orhan iyi+dir* = Orhan va bien », souligne la difficulté que pose ce suffixe pour des apprenants débutants : « *the suffix -Dir is a*

complicated element and it is probably best to postpone a detailed discussion of its uses » (p.32). Ici, *-Dir* est présenté comme un suffixe de la 3^e personne qui est employé lorsque le locuteur souhaite appuyer la valeur de vérité ou de généralisation de son assertion, comme dans l'exemple suivant :

(7)

<i>Atlar</i>	<i>tembel.</i>
cheval + pluriel	paresseux

Les chevaux sont paresseux.

<i>Atlar</i>	<i>tembeldir.</i>
cheval + pluriel	paresseux + dir

Tous les chevaux sont paresseux.

Selon l'auteur, le suffixe est *communément* employé à l'écrit, notamment dans la presse écrite, mais *beaucoup moins fréquemment* dans le discours oral dans un registre informel. Une fois de plus, les exemples fournis sont en décalage avec les contextes d'usage d'écrits : « Orhan va bien » ou « tous les chevaux sont paresseux » ne relèvent pas *a priori* de l'écrit journalistique.

Dans *Méthode de turc Volume 1* (Bozdémir, 2023), *-Dir* est introduit dès les premières leçons et apparaît parmi les « suffixes prédicatifs du verbe être » (p.25), entre parenthèses, suivi de la mention « facultatif » :

(y)im	je suis
-sin	tu es
(dir)(facultatif)	il est

Avec une référence à la fréquence de la valeur copule, il est précisé que *-Dir* est « facultatif et **généralement** ignoré, il n'est utilisé que pour souligner l'information et ceci quand l'information est un fait général que tout le monde connaît » (p.26), avec l'exemple suivant :

(8)

Güneş	sıcaktır.
soleil	chaud + tır

Le soleil est chaud.

Une certaine progression didactique émerge alors dans la description de *-Dir* dans la mesure où dans *Méthode de turc Volume 2* (Berk & Bozdémir, 2014), ouvrage destiné à des apprenants plus avancés, il est précisé que « -dir peut avoir un sens de supposition ou bien de certitude selon le contexte » (p.103). De plus, dans le chapitre « Discours politiques », les auteurs affirment que « -dir est facultatif, mais dans un discours officiel il est **souvent** utilisé » (p.197). Les valeurs modales et la variable « genre discursif » sont ainsi présentes, et on trouve une fois de plus une hypothèse quant à la fréquence de la forme dans un genre donné.

Dans *A Student Grammar of Turkish* (Ketrez, 2012), *-Dir* n'est jamais présenté comme copule : la 3^e personne du singulier ne porte aucune terminaison, même facultative, pour exprimer le verbe « être ». *-Dir* apparaît tardivement dans l'ouvrage et sa description porte sur la modalité :

-Dir can be attached to nouns and verbs and expresses two meanings. One is fact (gerçek). The other one is guess (tahmin). We can attribute these meanings to the sentences depending on their context of utterance. In some cases, when the context is not known, the sentences can be ambiguous. (Ketrez, 2012 : 113)

En plus du contexte, Ketrez fournit deux clés linguistiques pour déterminer la valeur du *-Dir* lorsqu'il est suffixé à certains temps et modes. Ainsi, lorsqu'il suit le présent *-lyor*, il exprime toujours la supposition ; lorsqu'il suit le présent duratif *-mAktA*⁴, il

⁴ L'auteure mentionne d'ailleurs que *-mAktA* exprime la même temporalité que *-lyor* mais est employé « dans des textes formels et des déclarations » (p.113). Cette caractéristique de *-mAktA* n'est pas évoquée dans les autres ouvrages, et constitue un indice supplémentaire quant à l'interprétation de *-Dir*.

exprime toujours un fait ; dans tous les autres cas, c'est bien le contexte qui doit guider l'interprétation. La combinaison *-mAktA+ Dlr* est « habituellement utilisée dans des contextes formels » (p.113).

L'ouvrage *Yedi İklim* (2021) introduit simultanément, dès le début du niveau B1, deux utilisations possibles de *-Dlr*, sous la dénomination globale de *pekiştirme* (emphase), avec (i) la valeur modale de supposition et (ii) le « langage officiel » ou l'« annonce » (p.14). Les critères proposés pour interpréter *-Dlr* relèvent donc du registre et du genre discursif, mais ces critères ne permettent pas de distinguer la valeur de supposition de celles liées à l'assertion. La différence écrit/oral n'est pas mentionnée. Dans les exemples liés à la supposition, certains exemples fournis comportent un adverbe ou une expression exprimant l'incertitude (9), tandis que dans d'autres, c'est le sens général de la phrase qui situe le propos dans un contexte d'incertitude (10) :

(9)

Fotoğraftaki	kişi	Oya'ya	benziyor.	Büyük bir	ihtimalle	annesidir.
photo + locatif + relatif	personne	oya + directif	ressembler + 3 ^e p.s.	grande	une	probabilité + avec mère + possessif 3 ^e p.s. + dir

La personne sur la photo ressemble à Oya. **C'est très probablement** sa mère.

(10)

Geç	kaldık.	Ders	çoktan	başlamıştır.
tard	rester+passé 1 ^e p.p.	cours	longtemps + ablatif	commencer + passé non constaté + tir

Nous sommes en retard. Le cours **a dû commencer** depuis longtemps.

Pour illustrer les usages liés aux « annonces », l'ouvrage propose des exemples tels que :

(11)

Yeni	dönem	kurs	kayıtlarımız	başlamıştır.
nouvelle	période	cours	inscription + pluriel + possessif 1 ^e p.p.	commencer + passé non constaté + tir

Les inscriptions pour la nouvelle période de cours ont commencé.

Ainsi, nous avons vu que dans plusieurs descriptions, au-delà de la simple fonction de copule, des valeurs modales parfois opposées sont attribuées à *-Dir*. Globalement, la valeur « incertitude » est plus souvent évoquée que la valeur « certitude ». Un seul ouvrage ne mentionne aucune valeur modale. Dans le tableau 3, nous reprenons les fonctions et valeurs modales attribuées au suffixe *-Dir* dans les ouvrages analysés.

	3e p. « être » (copule)	Supposition	Emphase, certitude ou vérité générale
Bazin	X	-	X
Aarsen & Bakus	X	-	-
Goldstein	-	X	X
Bozdémir	X	-	X
Berk & Bozdémir	-	X	X
Lewis	X	X	X
Underhill	X	-	X
Ketrez	-	X	X
Yedi İklim	-	X	X

Tableau 3. Fonctions et valeurs modales attribuées à *-Dir*

Pour identifier la valeur modale exprimée par *-Dir*, les auteurs proposent une série de variables ou d'indices, que nous reprenons dans le tableau 4.

	Contexte morphologique immédiat	Registre formel / informel	Écrit / oral	Genres de discours	Diachronie
Bazin	X		X		
Aarsen & Bakus					
Goldstein	X	X			
Bozdémir					
Berk & Bozdémir				Discours politiques	
Lewis	X	X	X		X
Underhill		X		Articles de journaux	
Ketrez	X	X			
Yedi İklim		X		Annonces	

Tableau 4. Les variables mentionnées dans les descriptions de *-Dir*

À travers ces tableaux, il apparaît que la complexité et la polyfonctionnalité de *-Dir* est davantage mise en valeur dans certains ouvrages. Parmi les différentes variables qui déterminent le sens de *-Dir* dans le discours, les plus consensuelles semblent être le contexte morphologique, c'est-à-dire le temps du verbe auquel *-Dir* est suffixé, et le registre. Globalement, pour un apprenant du turc L2 qui se trouve exposé à ces descriptions diverses, la compréhension de *-Dir* et son emploi adéquat restent à notre sens hasardeux. De plus, nous avons observé que les auteurs font régulièrement référence à la fréquence de telle ou telle fonction ou valeur modale de *-Dir* dans la langue. Pourtant, ces estimations de fréquence reposent probablement sur l'intuition des auteurs. Une analyse de corpus permettrait d'apporter des éclaircissements sur les fonctions et les valeurs modales de *-Dir* dans la langue et leur fréquence. Enfin, nous avons observé que certains exemples proposés par les auteurs ne correspondent pas aux contextes d'utilisation recommandés dans leurs descriptions. Les corpus sont susceptibles de combler ces lacunes. Un certain nombre de critères régulièrement évoqués dans les ouvrages (le contexte morphologique immédiat, le registre de langue, et parfois le genre discursif) ont retenu notre attention et orienté nos recherches sur le *TNC*.

3.2. La forme *vardır* et ses cooccurrences modales dans le *Turkish National Corpus v.3*

3.2.1. La forme *vardır* à l'écrit et à l'oral

Avant tout, nous avons recherché la fréquence relative de *vardır* dans la totalité du corpus écrit et oral du TNC. Afin d'évaluer si *vardır* est significativement plus fréquent à l'oral qu'à l'écrit dans ce corpus, un test du rapport de vraisemblance (*log-likelihood*) a été réalisé en comparant sa fréquence dans les deux sous-corpus. La forme apparaît 32 934 fois dans les données écrites et 759 fois dans les données orales. Bien que les fréquences relatives soient relativement proches (66,46 occurrences pour 100 000 mots à l'écrit contre 77,85 à l'oral), la différence est très significative ($p < 0.0001$) : la forme est surreprésentée dans le corpus oral. Ces résultats sont repris dans le tableau 5.

	Taille du corpus	Nombre d'occurrences	Fréquence par 100 000 mots
écrit	49 559 700	32 392	66,46
oral	974 857	759	77,85
total	50 534 557	33 693	

Test *log-likelihood* ($p < 0.0001$)

Tableau 5. Les occurrences de *vardır* dans le TNC

Si nous nous penchons maintenant sur les différents genres que nous avons choisi d'étudier dans ce corpus, nous observons que c'est dans le sous-corpus « écrit scientifique » que la forme *vardır* est la plus fréquente (115,1 occurrences pour 100 000 mots). La fréquence relative est plus faible dans le sous-corpus d'écrit non-scientifique (72,7) et *a fortiori* à l'écrit informel (40,7). Le test du *log-likelihood* appliqué aux trois sous-corpus révèle une différence très significative ($p < 0.0001$). Le tableau 6 reprend ces résultats.

	Taille du corpus	Nombre d'occurrences	Fréquence relative
Écrit scientifique	5 497 606	6 328	115,1
Écrit non scientifique	5 055 382	3 674	72,7
Écrit informel	1 667 965	678	40,7

Test *log-likelihood* ($p < 0.0001$)

Tableau 6. Les occurrences de *vardır* dans les sous-corpus écrits

S'agissant cette fois des genres oraux, la forme *vardır* est très fortement surreprésentée dans le sous-corpus « Assemblée nationale », avec une fréquence relative de 153,5 pour 100 000 mots, par rapport au sous-corpus « conversation » (fréquence relative de 32,6). Le tableau 7 reprend les résultats concernant les sous-corpus oraux.

	Taille du corpus	Nombre d'occurrences	Fréquence relative
Assemblée nationale	83 381	128	153,5
Conversation	426 443	139	32,6

Test *log-likelihood* ($p < 0.0001$)

Tableau 7. Les occurrences de *vardır* dans les sous-corpus oraux

Notre objectif à présent est de tenter de comprendre si la forme *vardır* exprime plutôt une incertitude/supposition ou une certitude/généralité. Nous regardons donc le contexte de 5 mots à gauche pour comparer les cooccurrences avec des adverbes modaux qui se trouvent sur une échelle incertitude/certitude.

Concernant l'écrit scientifique, nous n'avons trouvé aucune cooccurrence entre *vardır* et les marqueurs modaux. Cela peut être interprété de deux manières : soit les marqueurs modaux recherchés ne sont pas caractéristiques de ce genre ; soit la forme *vardır* n'a pas de valeur modale dans ce genre : le suffixe est alors simplement un

marqueur du registre formel, ou du langage soutenu, interprétation qui rejoint celle de Goldstein (1999).

Dans le sous-corpus d'écrit non scientifique, des cooccurrences de *vardır* ont été observées avec les marqueurs de certitude "*elbet(te)*" et "*mutlaka*", avec une mesure d'association élevée. Dans les exemples (12) et (13) tirés du corpus « écrit non scientifique », ces deux adverbes permettent d'interpréter le suffixe *-Dir* comme ayant une valeur de certitude et d'assertion forte :

(12)

<i>Şehrin</i>	<i>gelişmesinin</i>	<i>özgün</i>	<i>dinamikleri</i>	<i>elbette</i>	<i>vardır.</i>
ville + génitif	développer + possessif + génitif	propre	dynamique + pluriel + possessif	bien sûr	il y a + <i>dir</i>

Le développement de la ville **a bien sûr** ses dynamiques propres.

(13)

<i>Gördüklerimizin,</i>	<i>yaşadıklarımızın</i>	<i>bilimsel</i>	<i>bir</i>	<i>açıklaması</i>	<i>vardır</i>	<i>mutlaka.</i>
voir + relative + 1 ^e p.p. + génitif	vivre + relative + 1 ^e p.p. + génitif	scientifique	une	explication	il y a + <i>dir</i>	absolument

Il y a forcément une explication scientifique à ce que nous voyons et vivons.

S'agissant de l'écrit informel (forums, blogs), tous les adverbes modaux recherchés apparaissent dans le contexte immédiat de *vardır*, mais les résultats les plus significatifs apparaissent dans les mesures d'association liés aux marqueurs *belki* (peut-être) et *mutlaka* (absolument). Nous illustrons ces cooccurrences avec des marqueurs exprimant l'incertitude (exemple 14) et la certitude (15).

(14)

<i>Hemen</i>	<i>panikledim</i>	<i>belki</i>	<i> Radyasyon</i>	<i>vardır</i>	<i> diye.</i>
immédiatement	paniquer + passé 1 ^e p.s.	peut-être	radiation	il y a + <i>dir</i>	que

J'ai immédiatement paniqué en pensant **qu'il y avait peut-être** de la radiation.

(15)

<i>Her</i>	<i>liderin</i>	<i>her</i>	<i>şirketin</i>	<i>her</i>	<i>insanın</i>	<i>mutlaka</i>	<i>bir</i>	<i>öyküsü</i>	<i>vardır.</i>
chaque	leader + génitif	chaque	entreprise + génitif	chaque	personne + génitif	absolument	une	histoire + possessif 3 ^e p.s.	il y a + <i>dir</i>

Chaque leader, chaque entreprise, chaque personne **a forcément** une histoire propre.

En ce qui concerne l'oral, dans le sous-corpus des échanges parlementaires, que nous avons identifié comme relevant d'un registre plus formel, nous n'avons pas observé de cooccurrence. La même interprétation que pour l'écrit scientifique s'impose ici : il se peut que les marqueurs modaux recherchés ne soient pas caractéristiques de ce genre ; ou bien que la forme *vardır* n'ait pas de valeur modale dans ce genre.

Dans le sous-corpus des conversations, les cooccurrences avec les marqueurs *belki*, *her(h)alde* et *kesin(likle)* sont plus nombreuses, avec une mesure d'association élevée. Nous notons dans les exemples (16), (17) et (18) que la transcription reflète la prononciation informelle de certains mots plutôt que l'orthographe régulière. Les versions « rectifiées » sont proposées entre parenthèses.

(16)

<i>Fatoş'ta</i>	<i>vardır</i>	<i>belki</i>	<i>ona</i>	<i>bi (bir)</i>	<i>sor.</i>
fatoş+locatif	il y a + <i>dir</i>	peut-être	elle + directif	une	demander + impératif 2 ^e p.s.

Peut-être que Fatoş en **a**, demande-lui donc.

(17)

<i>Fazla</i>	<i>gelen</i>	<i>bi (bir)</i>	<i>valizi</i>	<i>de</i>	<i>en</i>	<i>aşşa (aşağı)</i>	20-30	<i>kilo</i>	<i>vardır</i>	<i>heralde (herhalde)</i>
trop	venir + relatif	une	valise + possessif 3 ^e p.s.	aussi	le plus	bas	20-30	kg	il y a + <i>dir</i>	probablement

Sa valise qui est de trop **fait** aussi **probablement** au moins 20-30 kg.

(18)

Vardır kesin	<i>bi (bir)</i>	<i>bakalım</i>	<i>mi ?</i>
il y a + <i>dir</i>	certainement un	regarder + optatif 1 ^e p.p.	particule interrogative

Il y en a **certainement**, on regarde ?

Les résultats concernant les cooccurrences dans l'ensemble des sous-corpus écrits et oraux sont repris dans le tableau 8. Pour chaque genre d'écrit, le nombre de cooccurrences (C) de *vardır* et des cinq marqueurs modaux est fournie dans la première colonne et la mesure d'association (LL) dans la deuxième colonne. Les mesures d'association significatives sont indiquées en gras.

	<i>belki</i>		<i>her(h)alde</i>		<i>kesin(likle)</i>		<i>elbet(te)</i>		<i>mutlaka</i>	
	C	LL ⁵	C	LL	C	LL	C	LL	C	LL
Écrit non scientifique	-		-		-		28	21,67	23	32,42
Ecrit informel	9	7,27	4	0,33	7	0,83	7	0,48	16	11,96
Conversations	10	35,14	7	15,81	4	19,99	-		1	2,05

Tableau 8. Les cooccurrences de *vardır* avec les marqueurs modaux

Les analyses de corpus ont révélé avant tout que la forme *vardır* est statistiquement plus fréquente à l'oral qu'à l'écrit. Parmi les genres écrits étudiés, c'est dans l'écrit scientifique que la forme est la plus fréquente. De même, parmi les genres oraux étudiés, la forme est plus fréquente dans les registres plus formels. La valeur modale ne peut être identifiée que dans certains genres. La forme *vardır* est plus fréquemment utilisée dans un contexte marqué par la certitude dans l'écrit non scientifique, tandis qu'elle est plus fréquemment employée dans un contexte marqué par l'incertitude dans l'oral informel.

Ces analyses, basées sur une forme unique, sont à compléter, notamment avec des analyses portant sur l'emploi de *-Dir* avec le temps présent (*yor+dur*), le présent duratif (*mAktA+Dir*), le passé non-constaté (*mlş+tlr*) et le futur (*CAk+tlr*).

⁵ Les mesures d'association indiquées dans le Tableau 8 sont issues de l'interface du TNC.

4. Perspectives didactiques

Nous suggérons ainsi que les corpus constituent une source d'information cruciale pour la description de certaines formes grammaticales ambiguës et que ces descriptions peuvent être utiles dans l'enseignement/apprentissage du turc L2. Cette étude constitue une occasion pour s'interroger sur l'exploitation pédagogique des corpus en classe de langue, avec l'objectif d'enseigner différemment certaines formes linguistiques de la langue cible. Nos propositions dans cette section sont formulées à partir des travaux dans le champ de l'exploitation pédagogique des corpus ; nous ne sommes pas en mesure d'en évaluer l'efficacité.

À partir du cas de *vardır*, nous proposons la démarche suivante pour aborder le suffixe *-Dir* en classe de langue. Il s'agit d'abord de déterminer à quel niveau cette forme devrait être abordée. En l'absence d'un référentiel du turc L2 correspondant aux niveaux du CECRL au-delà du niveau A1 (Polat & al., 2018), notre choix est guidé par les objectifs d'enseignement/apprentissage propices à l'introduction de *-Dir*, étant donné ses descriptions que nous avons exposées dans cette étude. En raison de la polyfonctionnalité de *-Dir*, des objectifs distincts devront être fixés à des niveaux distincts. Ainsi, nous considérons qu'une approche progressive est nécessaire : à partir du niveau B1, la présentation du suffixe serait axée sur l'expression de la supposition à l'oral ; au niveau B2, l'utilisation de *-Dir* dans l'écrit formel pourrait être abordée.

Étant donné que le TNC n'est actuellement pas accessible au grand public, c'est l'approche indirecte (Yoon & Jo, 2014) qui s'impose : l'enseignant devra faire une sélection et proposer aux apprenants d'observer les différents usages de *-Dir* dans des textes appartenant à des genres variés. À défaut de pouvoir accéder au TNC, la constitution d'un corpus représentatif d'un genre écrit est relativement facile et pourrait être réalisée par les apprenants : déclarations officielles, discours politiques, *posts* de forums... La forme choisie pour cette démarche pourrait être *vardır* ou encore une des combinaisons possibles du suffixe *-Dir* avec un verbe fréquent tel que « *yapmak* » (faire) : *yapıyordur* (présent) ; *yapmaktadır* (présent duratif ou formel) ; *yapmıştır* (passé non-constaté) ; *yapacaktır* (futur). Le nombre d'items qui sera fourni aux apprenants ne peut être fixé ; cependant, il s'agit de fournir un grand nombre d'occurrences, bien au-delà des quelques exemples habituellement proposés dans

les manuels. Gilquin (2021) propose ainsi une liste de 80 occurrences pour l'étude de certaines constructions causatives en anglais L2.

L'objectif concerne d'une part la réception, à savoir l'interprétation correcte de *-Dir* et l'identification des éléments contextuels qui pourraient confirmer cette interprétation, et d'autre part la production. Au niveau B1, on pourrait demander aux apprenants d'interpréter le degré de certitude exprimé dans les différentes occurrences. Une manière de systématiser l'emploi approprié de *-Dir* à l'oral serait de provoquer des suppositions par des questions auxquelles la réponse ne peut pas être connue avec certitude, éventuellement sous forme de jeu. Au niveau B2, la forme pourra être réemployée dans le cadre d'une production écrite, dans un registre formel. Éventuellement, les genres oraux relativement formels tels que des discours pourront également permettre une utilisation adaptée de *-Dir*. Des approches contrastives pourront également être adoptées, par exemple en demandant aux apprenants d'établir si *-Dir* est typiquement employé plutôt dans les phrases issues d'articles de journaux ou dans celles tirées de déclarations officielles.

Conclusion

D'un point de vue didactique, nous pensons, comme d'autres, que le suffixe *-Dir* ne devrait pas être introduit aux apprenants du turc L2 dans sa fonction de copule. La valeur modale de l'expression de l'incertitude à l'oral informel, puis la valeur modale de l'expression de l'assertion forte à l'écrit dans un registre formel pourront être enseignées progressivement à partir du niveau B1. Une approche par genres nous semble nécessaire : les analyses de corpus ont permis de confirmer que certains genres sont plus propices que d'autres à l'utilisation de la forme *vardır*. Une distinction très nette entre production orale et écrite est également nécessaire, puisque le suffixe, d'après nos analyses, n'a pas la même signification dans ces deux situations. Il nous semble utile que les apprenants puissent accéder à une quantité importante de données afin de découvrir intuitivement certains résultats que nous avons pu mettre en évidence à travers notre analyse de corpus. Nous faisons également écho aux travaux mettant en avant la nécessité de proposer aux enseignants et aux futurs enseignants de langue une formation initiale et continue à l'exploitation pédagogique des corpus.

Références bibliographiques

- Aarsen, J., & Backus, A. (2001). *Colloquial Turkish*. Routledge.
- Aksan, Y., et al. (2012). Construction of the Turkish National Corpus (TNC). *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC 2012)*. <http://www.lrec-conf.org/proceedings/lrec2012/papers.html>
- Aykurt-Buchwalter, S. (2023). Vers une « littératie de corpus » des enseignant(e)s de FLE : discours enseignants en contexte de formation continue. *Le Français dans le monde. Recherches et applications*, (74), 153-163.
- Bazin, L. (1997). *Introduction à l'étude pratique de la langue turque*. Maisonneuve.
- Berk, S. & Bozdémir, M. (2014). *Méthode de turc. Volume II. L'Asiathèque*.
- Boulton, A. (2024). Postface: Corpus et didactique en France. Virginie Privas-Bréauté. *Du recueil de données à l'analyse des corpus en didactique des langues*, Presses Universitaires de Rennes, pp 173- 178, 2024.
- Bozdémir, M. (2023). *Méthode de turc. Volume I. L'Asiathèque*.
- Cheng, W. (2010). What can a corpus tell us about language teaching? Dans A. O'Keeffe & M. McCarthy (Éds.), *The Routledge handbook of corpus linguistics* (p. 299–312). Taylor & Francis.
- Ekmekçi Aşan, G. (2021). Türk dilinin tarihi gelişimi içinde bildirme ekinin yapısal ve anlamsal işlevleri. *International Journal of Social and Humanities Sciences Research*, 8(74), 2440-2457. <https://doi.org/10.26450/jshsr.2684>
- Farr, F., & Leńko-Szymańska, A. (2024). Corpora in English language teacher education : Research, integration, and resources. *TESOL Quarterly*, 58(3), 1181-1192. <https://doi.org/10.1002/tesq.3281>
- Gilquin, G. (2021). Using corpora to foster L2 construction learning : A data-driven learning experiment. *International Journal of Applied Linguistics*, 31(2), 229-247. <https://doi.org/10.1111/ijal.12317>
- Götz, S., & Granger, S. (2024). Learner corpus research for pedagogical purposes : An overview and some research perspectives. *International Journal of Learner Corpus Research*, 10(1), 1-38.
- Johns, T. (1991). Should you be persuaded: Two examples of data-driven learning materials. *English Language Research Journal*, 4, 1-16.
- Ketrez, N. (2012). *A student grammar of Turkish*. Cambridge University Press.
- Korkmaz, Z. (2009). *Türkiye Türkçesi grameri. Şekil bilgisi*. Türk Dil Kurumu.
- Lewis, G. (2001). *Turkish grammar* (2e éd.). Oxford University Press.
- McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. Taylor & Francis.
- Mochizuki, M. & Aizawa, K. (2000). An affix acquisition order for EFL learners: An exploratory study. *System*, 28(2), 291-304. [https://doi.org/10.1016/S0346-251X\(00\)00013-0](https://doi.org/10.1016/S0346-251X(00)00013-0)
- Polat, Y. et al. (2018). *Yabancı Dil Olarak Türkçe İlk Kazanımlar İçin Başvuru Rehberi. A1 Öncesi*. Yunus Emre Enstitüsü
- Sripicharn, P. (2010). How can we prepare learners for using language corpora? Dans O'Keeffe & M. McCarthy (Éds.), *The Routledge handbook of corpus linguistics* (p. 371-384). Taylor & Francis.
- Tura Sansa, S. (1986). DIR in Modern Turkish. In A. Aksu-Koç & E. Taylan (Éds.), *Modern Studies in Turkish Linguistics: Proceedings of the 2nd International Conference on Turkish Linguistics*. Istanbul: Boğaziçi University Press.
- Underhill, R. (2003). *Turkish grammar*. The MIT Press.

Yoon, H., & Jo, J. W. (2014). Direct and Indirect Access to Corpora: An exploratory case study comparing students' error correction and learning strategy in L2 writing. *Language Learning and Technology*, 18(1), 96-117.

Zareva, A. (2017). Incorporating corpus literacy skills into TESOL teacher training. *ELT Journal*, 71(1), 69-79. <https://doi.org/10.1093/elt/ccw045>