

## La expresión de la probabilidad en el español no nativo

### Currículo y realidad

María José Barrios Sabador

Universidad Nebrija.

mjbarrio@nebrija.es

Barrios Sabador, M. J. (2013). La expresión de la probabilidad en el español no nativo: Currículo y realidad. *Revista Nebrija de Lingüística Aplicada* (2013) 13.

### RESUMEN

La expresión de la probabilidad es una de las áreas de la gramática española que provoca mayor número de errores entre los estudiantes extranjeros, incluso después de su enseñanza formal y su ejercitación. Este estudio se centra en el empleo de los operadores de probabilidad en estudiantes de nivel B2. Además acudimos a un grupo de nivel C1 como eje de comparación, así como a un grupo de hablantes nativos de español – grupo de control –. Las respuestas distintos tipos de actividad – todas ellas de carácter controlado – muestran una gran variabilidad y una falta de consistencia en los errores cometidos por los informantes.

Palabras clave: expresión de la probabilidad, B2, gramática, español segunda lengua

### ABSTRACT

*Expression of probability is an area of Spanish Grammar that causes a big rate of mistakes among foreign students, even after its formal instruction and practice. This research focuses on the use of probability components among students belonging to a B2 level of proficiency. We also consider a group of students from C1 level, who serve as comparison term, as well as a group of native Spanish speakers – control group –. The answers to different types of activity – all of them controlled – show evidence of the variability and lack of consistency in the errors made by informants.*

*Keywords: expression of probability, B2, grammar, Spanish Second Language*

## 1. OBJETIVOS DEL ESTUDIO Y COMPOSICIÓN DE LOS GRUPOS DE INVESTIGACIÓN

En nuestra labor en el aula, la expresión de la probabilidad es una de las cuestiones lingüísticas con mayor índice de errores entre los estudiantes, incluso después de su enseñanza formal y su reiterada ejercitación.

El objetivo principal de esta investigación fue observar si en el nivel B2 los informantes poseían el conocimiento de los recursos de expresión de la probabilidad marcados por el *Plan Curricular del Instituto Cervantes –PCIC–* para dicho nivel. Junto a él, nos marcamos un objetivo adicional, consistente en examinar si los alumnos que habían llegado ya a un nivel C1 dominaban dichos exponentes. Un conjunto de hablantes nativos no lingüistas nos sirvió como grupo de control. Básicamente, de forma sintética, nuestras hipótesis de investigación giraban en torno a las siguientes cuestiones:

- Qué operadores experimentaban un mayor número de errores y la posible relación entre tales dificultades y su condición modal: selección única (indicativo/ subjuntivo) o doble.
- Cuál era la influencia de los adverbios gradativos en la selección de indicativo en operadores como *lo más seguro es que, lo más probable es que oes muy probable que*.
- Qué errores producía la probabilidad morfológica.

- Cómo el tipo de actividad, según su grado de control, podía condicionar las respuestas a un mismo operador.

Dada su condición descriptiva, este estudio no es de naturaleza experimental, al no existir una manipulación deliberada de las variables. Nuestra *variable independiente* fue el nivel de competencia en la expresión de la probabilidad en español, preexistente a la realización de la prueba. La *variable dependiente* fue el conocimiento de la expresión de la probabilidad en español por parte de nuestros informantes. Asimismo, tuvimos cuatro *variables moderadoras*, representadas por los cuatro modelos de ejercicio.

Nuestro muestreo puede calificarse como probabilístico, por conglomerados, ya que nuestros informantes se asocian en una agrupación natural, al estar constituidos los grupos por alumnos de distintas clases de nuestro centro de trabajo. En la investigación participaron dos grupos con presencia de la variable independiente, formados por cuatro clases de nivel B2, con un total de treinta y dos informantes, y dos clases de nivel C1, con un total de catorce individuos. La media de edad de los sujetos, universitarios estadounidenses, era de veinte años, con un semestre de estancia en Madrid. Como elemento de comparación, se acudió a veintisiete hablantes nativos, que constituyeron el *grupo de control*. Ninguno de ellos tenía formación lingüística especializada ni se dedicaba a la enseñanza de lenguas.

Seguidamente, damos cuenta de las características del estudio y de la justificación de la tipología de actividades escogida.

---

## 2. DESCRIPCIÓN DEL ESTUDIO

---

Nuestra investigación consta de cuatro secciones, todas ellas consistentes en ejercicios controlados en distinta medida. Tal decisión obedece a dos razones. Por una parte, queríamos investigar si los informantes, dependiendo de su nivel de dominio, realizaban de forma más satisfactoria determinadas actividades. Por otra parte, pretendíamos comprobar si el grado de control de la actividad influía en los resultados para un mismo operador. Este último objetivo se apoyó en el fenómeno, ampliamente estudiado, de la variabilidad de los datos obtenidos – en los planos fonológico, morfológico y sintáctico – entre no nativos según el tipo de técnica de recolección empleada. Según Tarone (1983/ 1992), la capacidad interlingüística del no nativo se compone de un contínuum de estilos. En un extremo se ubica el *estilo cuidado*, donde se presta atención a la elección de las formas lingüísticas – ante la necesidad de ser ‘correcto’ –. En el otro extremo surge el *estilo vernáculo*, cuando se efectúa una selección espontánea de las formas lingüísticas, como ocurre en la conversación natural. Así, en técnicas que propicien un uso informal de la lengua aparecerán estructuras simplificadas, no relacionadas con la lengua materna ni con la lengua objeto. Por el contrario, en actividades centradas en un manejo cuidadoso de la lengua, la producción del alumno contendrá más formas de prestigio de la lengua materna y de la lengua objeto.

La elección de actividades controladas para nuestra prueba se debió a la necesidad de medir el *conocimiento explícito* del alumno, entendiendo como tal aquel sobre el cual los usuarios son plenamente conscientes, es decir, el conocimiento sobre el lenguaje y los usos que puede tener (Ellis 2004: 235).<sup>1</sup> La exclusión de cualquier producción espontánea como objeto de análisis residió en el hecho de que, a pesar de su indudable ventaja como muestra más auténtica de la interlengua del alumno, con menor influencia de monitorización, encierra el inconveniente de no garantizar la aparición de las estructuras objeto de análisis. Las actividades controladas, por el contrario, dirigen la atención hacia las formas objeto de investigación. Además, permiten experimentar con mayor número de sujetos y ofrecen mayor facilidad en su cuantificación. En cuanto a sus desventajas, adolecen de artificialidad, lo que puede conllevar problemas de validez interna (Chaudron 2003: 790). Para prevenir, en cierta medida, tal efecto adverso, se incluyó una diversificación en la tipología de ejercicios, con el fin de efectuar una triangulación de datos que otorgase mayor validez a los resultados.

En el primer ejercicio, que buscaba la producción espontánea semidirigida, se pedía al alumno que efectuara cuatro conjeturas sobre una situación extraña emplazada en el ámbito temporal del pasado. Para ello se le proporcionaron tres operadores, con cada uno de los cuales debía formar una oración (*quizás, a lo mejor, puede que*). Se dejó una cuarta línea en blanco, con el fin de comprobar si, ante la ausencia de operadores, el alumno recurría al morfema de condicional.

El segundo ejercicio consistió en un minidiálogo (compuesto de doce ítems, cuatro de ellos distractores) contextualizado, donde el alumno había de conjugar el verbo de una oración en el modo y tiempo adecuados. En esta actividad, unos amigos comentan la extraña reacción de un compañero. Incluimos dos ítems a modo de ejemplo:

1. Quizás aquel día (estar, él) \_\_\_\_\_ en el hospital por algún problema de salud.

2. ¡Qué negativo eres! Puede que ese fin de semana (irse, él) \_\_\_\_\_ de vacaciones.

El tercer ejercicio, formado por treinta y dos ítems, diez de ellos distractores, centraba su interés en la elección del operador apropiado para una oración, inserto en un minidiálogo. Sirva de muestra el ítem 7:

- Ayer por la tarde, Pedro tuvo problemas de estómago.
- \_\_\_\_\_ fue por la mayonesa, que le sentó mal.

- a)Igual  
b)Puede que  
c)Es muy probable que

Con el fin de observar la consistencia en la actuación del informante, juzgamos necesaria la utilización de cada operador en más de una ocasión a lo largo del ejercicio. Por ello, en él figuraban dos ítems cuya respuesta correcta respondía a un mismo operador. Para los operadores de doble selección (*quizás, posiblemente y seguramente*), uno de los ítems se hallaba con el verbo en modo indicativo y el otro con el verbo en subjuntivo.

Por último, el cuarto ejercicio pedía al informante que juzgase la idoneidad o aceptabilidad de una serie de oraciones y que, en el caso de juzgarlas incorrectas, las corrigiera, a la vez que señalaba su grado de seguridad en la respuesta. La actividad constaba de cuarenta ítems, once de los cuales eran distractores. Incluimos algunos como ejemplo:

	Correcta	Incorrecta	Tu corrección	En mi respuesta estoy:		
				Completa-mente seguro	Bastante seguro	Poco seguro
1. Posiblemente Rosa se fuera de casa antes de las diez.						
2. Pensaba ir al cine mañana.						
3. Lo más seguro es que ellos fueran los primeros en llegar al hotel.						
4. Es posible que esas fotos eran robadas.						

El diseño de la prueba planteó dificultades en cuanto a su extensión. Por una parte, dado el elevado número de exponentes de probabilidad que figura en el PCIC para el nivel B2, nos vimos obligados a efectuar una reducción que incluyera los de uso más común y que contemplara un equilibrio entre los que eligen indicativo, subjuntivo y los de doble selección modal. Así conseguiríamos que cada exponente pudiese aparecer más de una vez, lo que garantizaría una mayor consistencia de los resultados. Por otra parte, la posibilidad de expresar incertidumbre sobre acontecimientos pretéritos, presentes o prospectivos exigía, de querer contemplar todos estos ámbitos, una prueba de amplias dimensiones. Por ello, y tomando en consideración que el punto de inflexión entre los niveles B1 y B2 respecto a la expresión de la probabilidad reside en la introducción de los tiempos de pasado de subjuntivo y de los morfemas de futuro perfecto, condicional simple y condicional perfecto, optamos por circunscribir todas las actividades de la prueba al ámbito temporal del pasado.

DOBLE SELECCIÓN MODAL	SELECCIÓN DE INDICATIVO	SELECCIÓN DE SUBJUNTIVO	MORFEMAS DE PROBABILIDAD
QUIZÁ(S) POSIBLEMENTE	IGUAL A LO MEJOR	PUEDE (SER) QUE ES POSIBLE QUE	FUTURO PERFECTO CONDICIONAL SIMPLE
SEGURAMENTE	SEGURO QUE	LO MÁS SEGURO ES QUE LO MÁS PROBABLE ES QUE ES MUY PROBABLE QUE	CONDICIONAL PERFECTO

### 3. RESULTADOS

Dados los límites del presente artículo, nos limitaremos a la realización de un breve análisis de los datos obtenidos, organizado por actividades. Posteriormente, en la discusión de dichos datos, se propondrán algunas hipótesis explicativas sobre la etiología de los errores observados.

Empezamos comentando el Ejercicio 1, el de menor control de la prueba, que arroja un elevado número de errores por uso de indicativo *con puede que*. También *lo mejor* se distingue por un escaso índice de aciertos, aunque superior al experimentado por el anterior. El empleo del morfema de condicional es nulo, lo cual no significa su desconocimiento, según constatan las respuestas correctas del ítem correspondiente en el Ejercicio 4. Por consiguiente, no es descaminado aventurar que la tipología del presente ejercicio haya condicionado su ausencia, al no poder ofrecer en el enunciado ninguna indicación explícita para su producción.

El Ejercicio 2 exhibe un alto porcentaje de aciertos en los operadores que exigen o permiten el modo indicativo. Claro que este dato no tiene por qué significar que el aprendiz conoce el uso modal del operador, pues cabe la posibilidad de que acuda al modo conocido, al modo por defecto. En otras palabras, el mayor éxito en estos operadores podría estribar en el uso de las formas verbales más dominadas por el informante. En contraste con ellos, se verifica el exiguo índice de éxito de B2 en los operadores requeridores de subjuntivo: *lo más seguro es que*, *es posible que* y *puede (ser) que*. En cualquier caso, y al margen de los errores en que se ve implicada la necesaria utilización del subjuntivo, sobresale el condicional como el exponente con menor número de respuestas correctas, posiblemente debido a la inviabilidad de indicar explícitamente la necesidad de su empleo, al igual que se apuntó para el Ejercicio 1.

La nota más característica del Ejercicio 3 – de mayor grado de control que los dos previos, al tratarse de una actividad de selección múltiple – es la variable actuación de los grupos no nativos, sobre todo de B2, entre los dos ítems de un mismo operador. Sobresale el alto número de errores de B2 en los operadores que exigen o permiten el subjuntivo. Entre los primeros, es notable el acusado descenso de aciertos en los exponentes con elementos gradativos (*lo más seguro es que*, *lo más probable es que* y *es muy probable que*). En cuanto a los segundos, los fallos se incrementan cuando a *seguramente* le sigue subjuntivo y en el ítem donde *posiblemente* rige indicativo. Finalmente, dos operadores selectores de indicativo como son *igual* y *lo mejor* cosechan abundantes errores.<sup>2</sup>

El Ejercicio 4, al igual que los anteriores, muestra la lógica superioridad de los nativos sobre los informantes de C1 y la mayor habilidad de estos sobre los de B2. Ahora bien, contrariamente a nuestras expectativas, en algunos ítems se percibe en el grupo nativo una variabilidad inexistente en las secciones previas, inconsistente con la notable actuación observada en dichas secciones. Con todo, prevalece, en líneas generales, la competencia del nativo en el discernimiento de los ítems incorrectos, donde supera holgadamente al aprendiz de B2 y de C1.

---

## 4. DISCUSIÓN

---

El estudio global de la actuación de los distintos grupos de informantes a lo largo de la prueba revela, como era esperable, la superioridad del grupo de control sobre los informantes no nativos. Las mayores inconsistencias surgen en el nivel inferior, el B2, a lo largo de las distintas secciones. El grupo C1 muestra, como se auguraba en un principio, mejores resultados y mayor coherencia en las respuestas. La mayor consistencia se percibe en el grupo nativo, aunque con ciertas salvedades que señalaremos oportunamente.

La tipología de errores es sumamente amplia, por lo que limitaremos nuestro análisis a los datos más recurrentes o sobresalientes. En líneas generales, el grupo B2 se caracteriza por una variedad de respuestas erróneas concernientes tanto al modo como al tiempo verbal seleccionado en los dos primeros ejercicios. Por su parte, los informantes de C1 ofrecen, en dichas actividades, un menor número de errores de carácter modal.

Globalmente, la complejidad en la expresión de la probabilidad se detecta en operadores de selección única – tanto de indicativo como de subjuntivo – y de doble selección modal. Los morfemas de probabilidad también originan una notable confusión, sobre todo en el grupo B2.

Entre los operadores selectores de indicativo, cosechan abundantes errores *seguro que*, *a lo mejor* e *igual*. En lo que concierne *a igual* y *a lo mejor*, una explicación a su dificultad ha de tener en cuenta su escasa transparencia semántica, al ser exponentes que poseen un significado claramente indescifrable desde el análisis de los elementos que los componen (los adjetivos/ adverbios *igual* y *mejor*). Es posible que este significado tan idiosincrásico sea responsable del cuantioso número de errores y del elevado coste de procesamiento que supone su comprensión y, por ende, su adquisición.

No es aplicable el mismo criterio explicativo *aseguro que*, operador de significado transparente, normalmente conocido por el estudiante de este nivel. Es posible que el contacto con otras opciones de respuesta en el Ejercicio 3 (como *lo más seguro es que* o *es muy probable que*) provoque su desestimación, al concebirse las otras alternativas como vehículo de mayor certidumbre. La presencia de la conjunción *que* puede ser asimismo responsable de la inapropiada recurrencia al subjuntivo.<sup>3</sup> Tampoco cabe excluir la eventualidad de que estemos ante una hipergeneralización de reglas: la exigencia de subjuntivo por parte de algunos operadores puede conducir al alumno a la selección del subjuntivo como modo por defecto, no marcado, para la expresión de la probabilidad. A ello contribuiría, desde otro plano, la común asunción por gran parte del alumnado del subjuntivo como transmisor de duda e incertidumbre.

Igualmente, operadores de uso exclusivo de subjuntivo causan un elevado número de contestaciones incorrectas. Entre estos últimos, son aquellos con presencia de elementos gradativos (*lo más seguro es que*, *lo más probable es que*, *es muy probable que*) los que ocasionan mayor confusión al aprendiz de B2 y de C1. Es posible que la adición del cuantificador induzca a la comprensión del operador como indicador de mayor certidumbre y, dada la asimilación 'indicativo – certitud' entre muchos estudiantes, lleve a la falsa inferencia de que requiere este modo. En el ámbito nativo, es destacable la inconsistencia de los juicios de gramaticalidad en el análisis de los ítems con *lo más seguro es que*, frente al índice de aciertos del operador en los ejercicios 2 y 3. Una hipótesis explicativa de este cambio de actuación es que en el Ejercicio 4 – juicios de gramaticalidad – surja en la mente del nativo cierto paralelismo entre el operador y la existencia de estructuras ecuacionales donde la nominalización del adjetivo abre las puertas a la doble selección modal:<sup>4</sup> *Lo más importante es que haya/ ha venido*.

Sin alcanzar el mismo nivel de respuestas erróneas de estos últimos, operadores como *puede que yes posible que* causan asimismo notables dificultades, a pesar de corresponder ambos al nivel B1 del PCIC.<sup>5</sup> Como explicación de tales resultados, juzgamos que la deficiente interpretación de las claves contextuales (presencia de justificaciones de la conjetura que llevarían a la falsa inferencia de que se comunica algo presumiblemente “cierto” o, en sentido contrario, la aparición de un *no sé* como desencadenante de un marco contextual de elevada incertidumbre) sería la responsable de la elección del operador o del modo incorrectos, pues son muchos los estudiantes que adjudican al modo por sí solo el valor semántico de certeza (indicativo) o inseguridad (subjuntivo), al margen del entorno discursivo en que se ubique.

Respecto a los operadores de doble uso modal *seguramente, posiblemente y quizá(s)*, se comprueba que inducen, igualmente, a un buen número de desaciertos e inconsistencias. Aunque en un principio se pensó que la viabilidad de empleo de ambos modos desembocaría en un uso correcto de estos operadores, los resultados demostraron más bien el fenómeno contrario. Una explicación plausible a tales dificultades radicaría en que la doble opción modal lleva al alumno a otorgar al modo un valor no distintivo, redundante y, como tal, poco rentable a efectos de aprendizaje. A ello se sumaría la existencia de reglas descriptivas nada claras en cuanto a la repercusión significativa de la elección de indicativo o subjuntivo. En cuanto a los razonamientos de tipo pragmático – discursivo que razonan esta doble posibilidad, resultan igualmente poco esclarecedores para el aprendiz, al residir en multitud de factores contextuales inaprehensibles y de compleja sistematización.

La revisión de los resultados desvela que *seguramente* es concebido en buena medida como operador de selección única de indicativo. Este adverbio, procedente del adjetivo *seguro*, definido como “Cierto, indubitable y en cierta manera infalible” según la RAE, es asociado por un buen número de aprendices al empleo exclusivo de indicativo, modo, en concepción de muchos, inviable para la expresión de incertidumbre. Tal creencia no surge entre los nativos, con la notoria excepción del Ejercicio 4, donde se discierne cierto rechazo (30%) al empleo de subjuntivo con este operador, frente a su plena aceptación en el Ejercicio 3.

Por último, la probabilidad morfológica es responsable de un elevado número de errores en B2. Estos resultados no son insospechados, dada la *complejidad funcional* de estos morfemas. Pensemos, por ejemplo, en los “valores” del condicional: morfema indicativo de acción prospectiva medida desde un punto temporal ubicado en el pasado, marcador de incertidumbre en un tiempo pretérito, de valor contrafactual en el presente de la enunciación o referente a acciones hipotéticas futuras. Por una parte, cabría prever que el valor *significativo* de la morfología de los futuros y condicionales, frente a sus correspondientes formas asertivas, facilitase el reconocimiento de estas formas, por lo que este rasgo debería ser suficiente para caracterizarlos como formas de alto *valor comunicativo*, en terminología de Van Patten (2002). Sin embargo, no parece que ello resulte suficientemente obvio para el aprendiz. Tal arduidad se vería acrecentada por su condición de formas *semánticamente complejas*, al ligarse sus reglas o condiciones de uso a factores discursivos y semánticos no fáciles de captar por el estudiante extranjero (Doughty y Williams 1998: 221 – 223). Basta tener presente la amplia gama de sentidos contextuales de futuros y condicionales (probabilidad, concesión, sorpresa...), intrínsecamente ligados a su entorno semántico y pragmático e ilustrativos tanto de su *complejidad funcional* como *significativa*.

---

## 5 CONCLUSIÓN

---

Los datos obtenidos en la prueba muestran una amplia gama de dificultades en el empleo de los operadores de probabilidad, no siempre dependientes del mayor o menor control del tipo de ejercicio y con porcentajes muy dispares en un mismo operador entre unas actividades y otras. No es factible justificar tan nutrido grupo de errores mediante un único argumento sin caer en la simplificación. Ahora bien, sí cabe afirmar que el amplio repertorio de operadores de que dispone el español para manifestar la incertidumbre, la diversidad modal que admiten, así como la presencia de morfemas verbales transmisores de distintos contenidos modales y temporales son, en grandes rasgos, factores que dificultan la adquisición de este aspecto gramatical por parte de los estudiantes de español. El camino que nos queda por recorrer es largo y no exento de dificultades, pero la labor del profesor siempre ha sido y será gratificante, también en el desafío que implica la explicación y didactización de este aspecto de nuestra gramática.

[1] A diferencia del explícito, el *conocimiento implícito* integra una serie de procesos de tipo intuitivo, de rápida ejecución, que requieren poca o ninguna atención. Este conocimiento se caracteriza por su automaticidad, frente al explícito, que implica, normalmente, procesos de tipo controlado.



[2] LaNGRALE (2009: 1956) señala el doble empleo modal de *lo mejor*, pese a la mayor extensión del indicativo, siendo la proporción de uso de subjuntivo ligeramente mayor en el español americano. Pese a ello, en nuestra prueba optamos por considerar tal elección como incorrecta, debido a nuestra variante diatópica: el español peninsular.

[3] En nuestra experiencia docente, es frecuente que el estudiante asuma, erróneamente, que la existencia de *que* condiciona la presencia de subjuntivo.

[4] Nótese el carácter factivo de tales formaciones, frente a la índole no factiva de *lo más seguro es que*.

[5] Aunque en B1 se ejemplifican únicamente con la forma de presente de subjuntivo.

## REFERENCIAS BIBLIOGRÁFICAS

AA. VV. (2006). *Plan curricular del Instituto Cervantes. Niveles de referencia para el español*. B. Madrid: Biblioteca Nueva.

Chaudron, C. (2003). "Data Collection in SLA Research". En C. Doughty y M. Long (Eds.), *The Handbook of Second Language Acquisition* (pp. 762 –828). Blackwell.

Doughty, C. y Williams, J. (1998). "Pedagogical choices in focus on form". En Doughty, C. y Williams, J (Eds.), *Focus on Form in Classroom Second Language Acquisition* (pp.197-261). Cambridge: Cambridge University Press.

Ellis, R. (2004). "The definition and Measurement of L2 Explicit Knowledge". *Language Learning*, 54 (2), 227 – 275.

Tarone, E. (1983/ 1992). "Dialectos idiosincrásicos y análisis de errores". En J. Muñoz Licerias (Ed.), *La adquisición de las lenguas extranjeras* (pp. 263 – 292). Madrid: Visor.

Real Academia Española. (2001). *Diccionario de la Lengua Española*. Madrid: Espasa Calpe.

<http://www.rae.es/rae.html>